Technical Note

Landslides
DOI 10.1007/s10346-024-02211-1
Received: 14 April 2023
Accepted: 3 January 2024

Benjamin Shafer · Beena Ajmera · Kamal Raj Upadhaya · Aaron Lee M. Daigh

Assessment of factors leading to the failure of slopes in North Dakota

Abstract A total of 66,894 landslides were observed in North Dakota. The characteristics of these landslide locations were compared with the properties of areas without landslides to assess the factors that may be contributing to the landslides. Specifically, 68,395 control point locations randomly distributed across the state were selected for these comparisons. All the landslides for this study were found in areas with slopes less than 64°, with the majority of the failures occurring on slopes with inclinations between 9° and 14°. The largest fraction of the landslides occurred in the Sentinel Butte Formation (34,063 or 51% of the total), followed by Bullion Creek (8695 or 13% of the total) and river sediment of the Oahe Formation (6421 or 9.6% of the total). In the *t* tests, all of the surficial geologic formations had statistically significant differences between the landslides and control points. The *t* test for the slope inclination indicated statistically significant differences with a *p*-value less than 0.001 and a huge effect size between the landslide and control points. The sodium adsorption ratio and total dissolved solids were also found to be statistically significant from the *t* test results. Pearson's correlation matrix showed a negative correlation between the amount of rainfall and various measures of the salt concentrations at the landslide locations, pointing to the reductions in shear strength and slope stability that might result as pore fluid salinity is leached.

Keywords Landslides · Statistical analysis · Correlation matrix · *t* tests · North Dakota

Introduction

In many regions across the world, landslides are frequent and lead to the loss of lives, infrastructure damage, large losses of money, and adverse environmental effects (Highland and Bobrowsky 2008). Each year over the last 20 years, landslides have had an average global impact on approximately 238,000 people, led to the deaths of 1000 people (1.5% of total deaths caused by natural disasters), and injured another 200 people (Ritchie and Roser 2014). During that time, the overall average economic damages due to landslides were recorded at $310 million per year (EMDAT 2020). The damages due to landslides in the USA vary from $2 to $4 billion, with 25–50 people losing their lives annually (Schuster 1978).

Radbrunch-Hall et al. (1982) stated that due to the low relief in the study area, landslides are uncommon in North Dakota. Their overview map of landslides in the continental USA indicated a low landslide incidence rate, which corresponds to landslides in less than 1.5% of the area across the state, with moderate susceptibility in the western parts of the state. Godt et al. (2012) and Mirus et al. (2020) also presented concurring maps. The existing models were noted to poorly capture the occurrence of landslides in terrains of low to moderate susceptibility (Mirus et al. 2020). One such terrain is North Dakota.

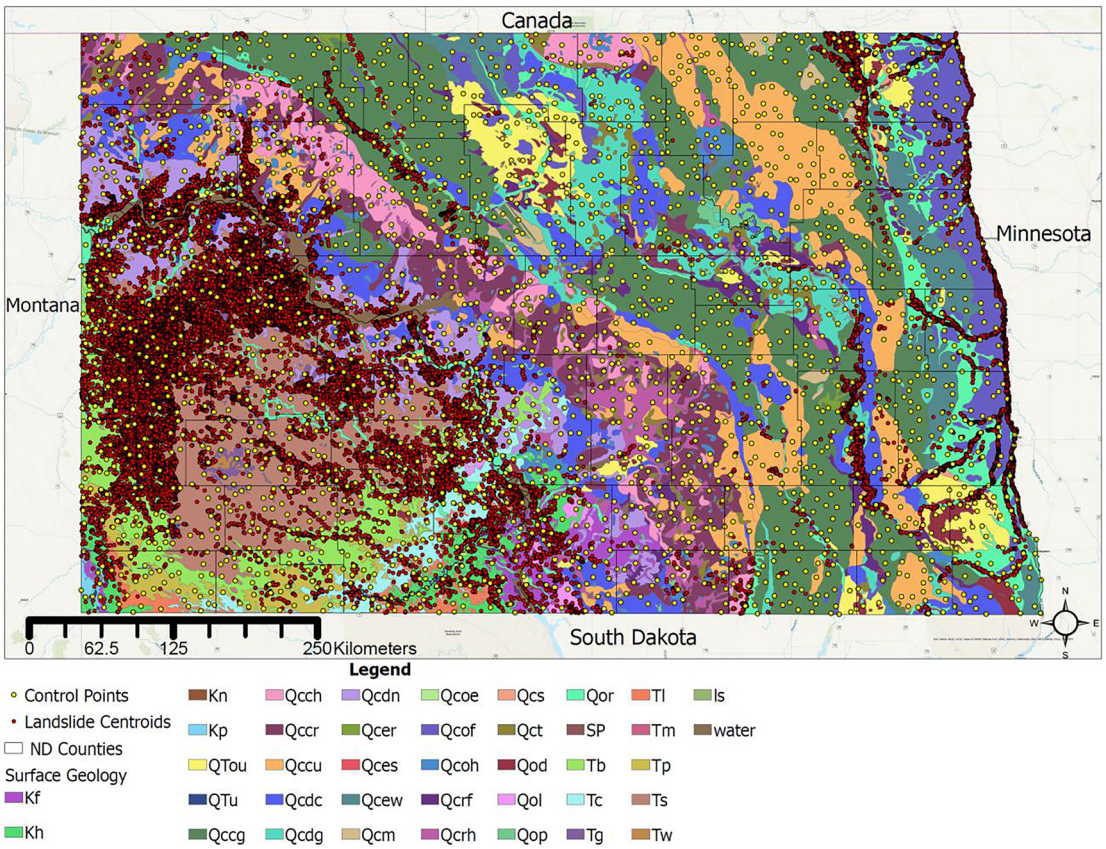
North Dakota is the site of thousands of landslides, which result in millions of dollars in damages each year (Murphy 2017). The results of these landslides are damage to infrastructure and an increased risk of flooding. Surficial geology will play a key role in the occurrence of landslides in the gently sloped terrain of North Dakota. However, this is not well captured through the coarse-grained geological mapping utilized by Radbrunch-Hall et al. (1982) or in the simple slope relief threshold used by Godt et al. (2012). In addition, several other factors, such as salinity, land cover, and land use, will also play a role in the occurrence of landslides across the state of North Dakota. Therefore, it is important to understand the conditioning factors that lead to these slope failures in order to reduce the damages they cause and prevent the occurrence of fatalities in the future. This study performed an analysis of the conditioning factors that lead to slope instabilities. This analysis is based on 66,894 historical landslides mapped and 68,395 control point locations randomly generated across the state of North Dakota. The results provide insights into the landslide distribution patterns relative to the geological and topographical characteristics, the size of the landslides, pore fluid chemistry, and weather patterns.

Study area

North Dakota is in the north-central portion of the USA, sharing its northern border with the Canadian provinces of Saskatchewan and Manitoba. The Red River makes up the eastern border of the state. Montana and South Dakota enclose the state along the western and southern borders, respectively. Geographically, the study area spans from 45°56′ N to 49°00′ N and 96°104′ W to 104°03′ W. The total area of North Dakota is approximately 183,121 km^2, with a majority of the state being land and only 2% of the area being composed of water (City-Data 2020). North Dakota approximately takes the shape of a geometric rectangle with maximum dimensions of 580 km E-W and 340 km N-S. The elevation of North Dakota ranges from 229 m along the eastern border to 1069 m in the southwestern part of the state.

Geology

The State of North Dakota encompasses 36 different surficial geologic formations (Clayton et al. 1980a). The five formations that cover the largest percentage of the state are collapsed glacial sediments, the Sentinel Butte Formation, Bullion Creek Formation, river sediments, and offshore sediments. A map of all the surficial geologic formations is included in Fig. 1. The characteristics of the three formations most relevant to this paper will be discussed in this section. Details for the remaining surficial geologic formations in North Dakota will not be included in this paper. For those interested, complete descriptions of all bedrock geologic formations are in Murphy et al. (2009) and in Clayton et al. (1980a) for surficial geological formations.



Legend	Formation Name	Legend	Formation Name
Kf	Fox Hills Formation	Kh	Hell Creek Formation
Kn	Undivided Niobrara and Carlile Formations	Kp	Pierre Formation
QTou	Sand of the Undivided Oahe and Older Formations	QTu	Undivided Quaternary and Upper Tertiary Sediment
Qccg	Gently Undulating Collapsed Glacial Sediment	Qcch	Hilly Collapsed Glacial Sediment
Qccr	Rolling Collapsed Glacial Sediment	Qccu	Undulating Collapsed Glacial Sediment
Qcdc	Collapsed/Draped Transition Sediments	Qcdg	Glacial Sediment Draped Over Pre-Existing Glacial Topography
Qcdn	Glacial Sediment Draped Over Pre-Existing Non-Glacial Topography	Qcer	River-Eroded Glacial Sediment
Qces	Slopewash-Eroded Glacial Sediment	Qcew	Wave-Eroded Glacial Sediment
Qcm	Glacial Sediment on Subglacially Molded Surfaces	Qcoe	Eroded Lake Sediment
Qcof	Proglacial-Lake Sediment	Qcoh	Ice-Walled-Lake Sediment or Collapsed Supraglacial-Lake Sediment
Qcrf	Uncollapsed River Sediment	Qcrh	Collapsed River Sediment
Qcs	Coleharbor Formation Shoreline Sediment	Qct	Glacial Sediment on Thrust Masses
Qod	Oahe Formation Windblown Sand	Qol	Oahe Formation Windblown Silt
Qop	Oahe Formation Pond Sediment	Qor	Oahe Formation River Sediment
Tb	Bullion Creek Formation	Tc	Cannonball Formation
Tg	Golden Valley Formation	Tl	Ludlow Formation
Tm	Undivided Upper and Middle Tertiary Rock	Tp	Slope Formation
Ts	Sentinel Butte Formation	Tw	White River Group

Fig. 1 Geological Map of North Dakota showing the locations of the centroids of the landslides and the control points. Base map from Clayton et al. (1980) and Bluemle (1983)

The Pierre Formation covers a total area of 939 km², which is approximately 0.5% of the study area. The composition of the Pierre Formation generally includes non-calcareous light to dark gray shale with fissile to blocky structures. The average depth of this formation is around 701 m (Murphy et al. 2009). The shale, in the Pierre Formation, comes from soils deposited in an offshore marine environment during the Cretaceous Era (Gill and Cobban 1965).

The Sentinel Butte Formation spans 22,965 km² (about 12.5%) of the land surface in North Dakota. It has a composition of a mix of grayish-brown clay, silt, sand, sandstone, and lignite. The average depth of the formation is about 200 m (Clayton 1980b). The steep slopes of the Badlands topography show evidence of tuffaceous beds, petrified wood, and Black Butte lignite. All of these are markers of the base of the Sentinel Butte Formation. Swelling bentonites, non-swelling claystone, limestone, poorly to well-cemented sandstone, and iron oxide are common in this formation (Murphy et al. 2009). The soils are generally from alluvial, lacustrine, and swamp depositional environments from the Paleocene Era (Jacob 1976).

The Bullion Creek Formation spans an area of 10,494 km². This is approximately 5.7% of the study area. The composition of the Bullion Creek Formation is a mix of yellowish-brown clay, silt, sand, sandstone, and lignite. The average depth of the formation is about 200 m (Clayton 1980b). The soil in the Bullion Creek Formation is usually classified as swelling and non-swelling claystones, limestone, poorly to well-cemented sandstone, and iron oxide (Murphy et al. 2009). The formation consists of soils deposited during the Paleocene Era in lacustrine, fluvial, and paludal environments (Clayton 1980b).

Climate

The variable climate of North Dakota is correlated with its position within the North American continent. The temperature varies dramatically throughout the year, reaching a maximum of 48 °C in the summer months and dropping as low as − 46 °C in the winter. The state experiences sporadic precipitation, low humidity, continuous wind, and consistent sunshine (NOAA 2017). North Dakota mainly receives cold, dry air from the north and hot, moist air from the south. This, in conjunction with the lack of cold, moist air from the west due to the mountains, leads to a variable climate (NOAA 2017). The study area is also prone to blizzards in the winter and tornados in the summer months (ACIS 2020). The start of the summer months in North Dakota can vary, but it is generally between July and August.

Based on ERA5-Land monthly averaged data for the years between 1979 and 2019 (Muñoz Sabater 2019), the total precipitation in depths of liquid water is presented in Figs. 2 and 3. Figure 2 shows that the average yearly precipitation fluctuates; it is typically between 35 and 75 cm each year. The average annual rainfall for the entire state ranges from 33 to 55 cm (ND 2020). The average snowfall for the entire state is 63 to 114 cm. The winter and snowfall months tend to land from November to March (NOAA 2017). Figure 3 shows the average monthly total precipitation over the same time span as Fig. 2. The majority of precipitation tends to occur between April and October.

Topography

North Dakota is primarily divided into two unique topographic regions. The first region is the Central Lowlands, which compose the northeastern part of the state. The second region is the Great Plains in the southwestern corner of the state (Bluemle and Biek 2007). The Central Lowlands extend westward from the Red River Valley, about 40 to 64 km and include the Missouri Escarpment, Turtle Mountains, Glaciated Plains, Pembina Escarpment, Souris Lake Plain, and Devils

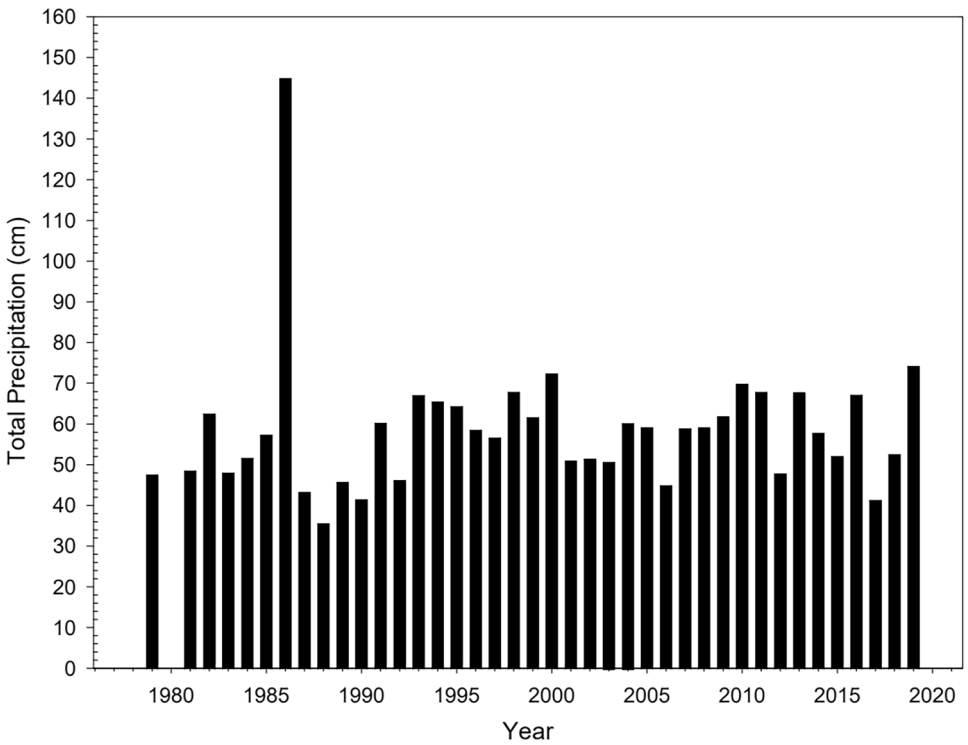


Fig. 2 Average annual total precipitation in North Dakota from 1979 to 2019 (Muñoz Sabater 2019)

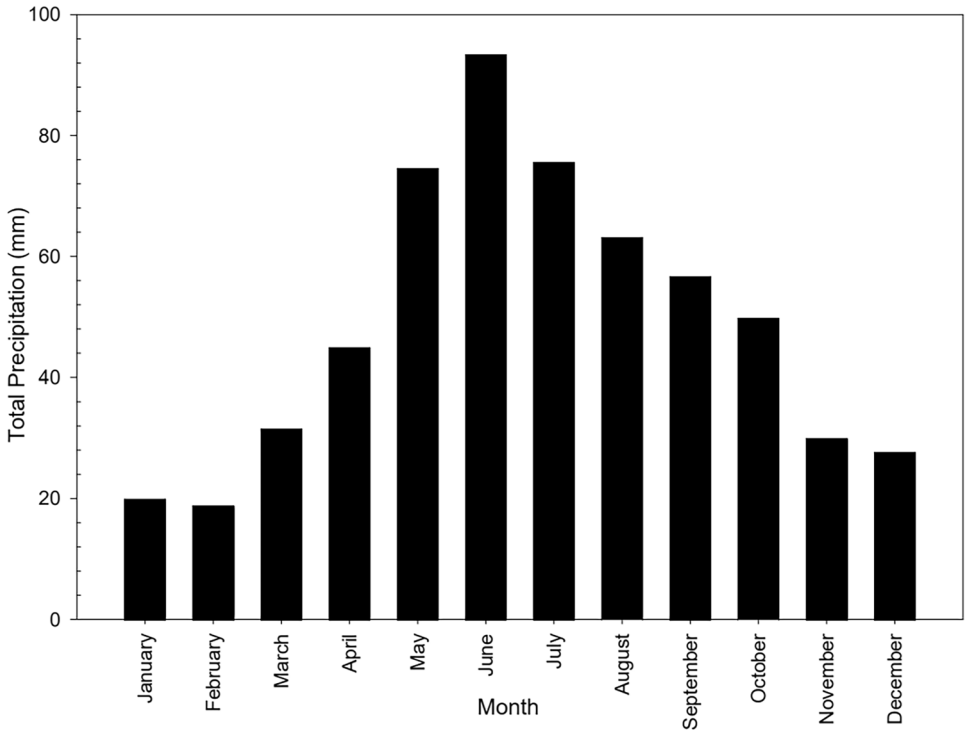


Fig. 3 Average monthly precipitation in North Dakota from 1979 to 2019 (Muñoz Sabater 2019)

Lake Basin. This expanse of land is relatively flat, with very few steep slopes. Thus, minimal landslides are reported in this area. The Great Plains in the southwestern portion of the state include plains of rolling hills along with the North Dakota Badlands. The Badlands are a more unstable topographic feature relative to the flatter topography of the eastern portion of the state due to their erosional faces and steep slopes (NOAA 2017). The Great Plains can be divided into the Missouri Plateau, the Little Missouri Badlands, the Coteau Slope, and the Missouri Coteau (Bluemle and Biek 2007).

The zone that makes up the center of the state acts as a transition between the two topographic regions and is called the Missouri Coteau. It extends from the northwestern corner to the south-central portion of the state (Bluemle and Biek 2007). The region has a measured elevation change from 91 to 152 m due to abnormal and glaciated landscapes. The abnormal topography in the area results from the collapse of superglacial soils. This led to a region with erosional bedrock covered in glacial deposits and plain surfaces. The glacial deposits on top of bedrock are more unstable and prone to failure than the plain surfaces (Bluemle and Biek 2007).

Data collection and analysis

The landslide locations used in this study were identified by the North Dakota Geologic Survey (NDGS). Since the authors did not map the landslides, interested readers are directed to Moxness (2019) for details of the mapping process. Only a brief summary is provided here due to space limitations. The identification of landslides in North Dakota started in the 1930s, when the United States Department of Agriculture (USDA) took aerial photos of the landslides using low-flying aircraft. The aerial photos were, then, viewed via 1:20,000-scale photographs in stereopairs. The official mapping of the landslides started in 2003 with the use of digital aerial orthophotography. The resolution has increased from 2 m in 2019 to 60 cm in 2018. By the beginning of 2017, 25% of the state area had been mapped (Murphy 2017). From 2017 through the start of 2019, the mapping effort had a rapid increase due to the availability of Light Detection and Ranging (LiDAR) data from the North Dakota State Water Commission (Anderson and Maike 2017; Moxness 2019). In these 2 years, a total of about 49% of the state had been mapped (Moxness 2019). The remaining approximately 51% of the state, mostly in the central lowlands of the eastern half of the state (Moxness 2019), was mapped between January 2019 and November 2021 (Moxness 2022). Figure 4 shows the distribution of the mapped landslides across the state of North Dakota.

The mapping efforts performed by the NDGS resulted in shape files for each identified earth flow, slump, and area of soil creep (Moxness 2019). These shape files, which included important information on the landslides, were used for further analyses using ArcGIS. All the data related to the conditioning factors were gathered via one of three methods: (i) based on the centroid of the mapped landslide, (ii) the largest percentage of the landslide area, or (iii) averaged over the area of the landslide. The control point locations were generated using the Create Random Points Tool in ArcGIS. Using the Select by Location tool, all control points that fell within a body of water, the boundary of a mapped landslide, or outside the study area were removed from the dataset. The locations of both the landslide and control point datasets are shown in Fig. 4 on a slope inclination map. A summary of all the datasets used for the conditioning factors is provided below:

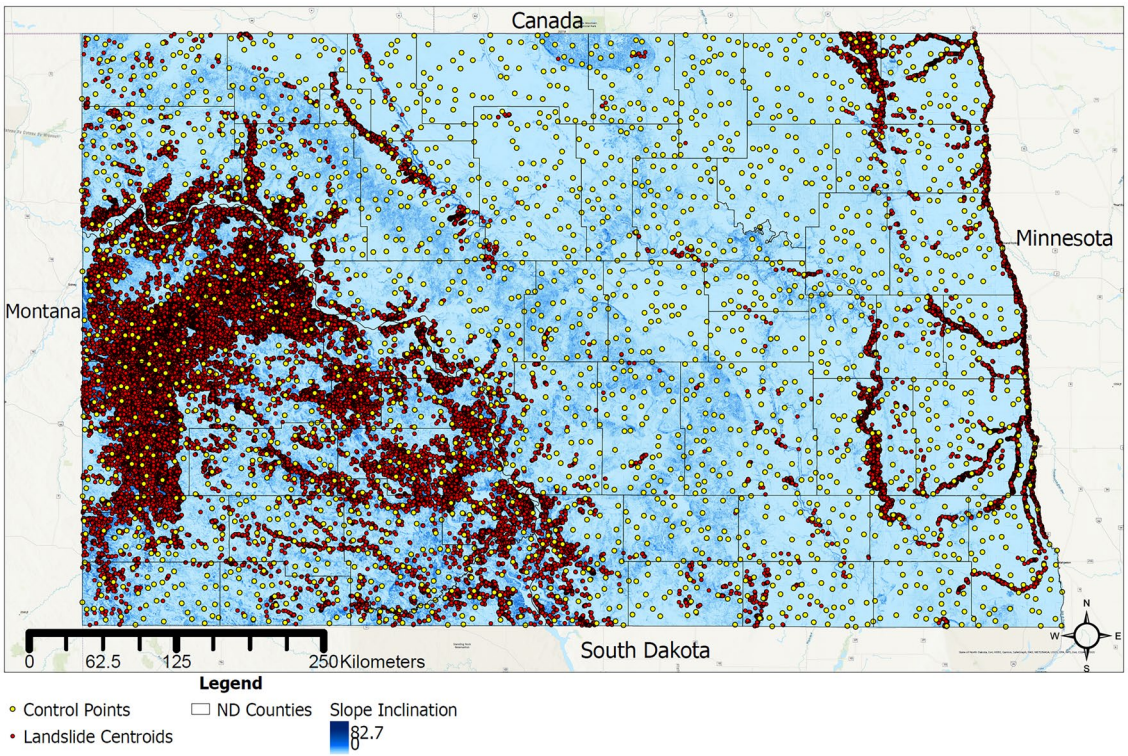


Fig. 4 Slope map of North Dakota showing the locations of the centroids of the landslides and control points. Slope inclination is presented in degrees

- Geologic formation (Fig. 1): The data were obtained from Clayton et al. (1980a) and Bluemle (1983) at a resolution of 250 m. Formations are described in detail in Clayton et al. (1980a). For this study, the assigned value for each landslide corresponded to the largest area percentage within its boundary.
- Slope inclinations (α) (Fig. 4): The data, with a resolution of 30 m, were obtained from the National Elevation Dataset (USGS 2019). Spatial analysis tools in ArcGIS were used to convert the elevation data into slope inclination. It is noted that the state, especially in the eastern regions, does not have major differences in slope inclination. The slope inclination at the centroid of the mapped landslide was assigned for the analyses presented in this paper.
- Land cover: The data were obtained from the National Land Cover Database (USGS 2016). This dataset had a resolution of 30 m. Land cover data from over many years were compared with minimal changes in the data. Therefore, the data from 2016 were used to select the value that corresponded to the largest percentage of the landslide area.
- Rainfall and snowfall: The data were obtained from NOAA (2020) at a resolution of 1 km. An average of 10 years of precipitation data were used since the time of failure of each landslide is not known. The rainfall and snowfall values were averaged across the area of the landslide.
- Electrical conductivity (EC) and sodium absorption ratio (SAR): The data were obtained from the Soil Survey Geographic Database (USDA 2013 and 2020). The resolution of this database is 25 m. EC and SAR values were averaged across the landslide boundary. Additional details are provided later in the text.

Numerous other conditioning factors may impact the stability of slopes in North Dakota, including the ecological region, glaciation, elevation, aspect, frost depth, amount of snow cover, etc. The study of the impact of these factors was beyond the scope of this paper. Rather, this paper focuses on the parameters listed above since they are expected to have the greatest impact on the stability of slopes in the study area.

Along with the data collection and analyses described previously, site reconnaissance for four landslides in western North Dakota was performed as part of this study. Object IDs with the corresponding landslide locations are presented in Table 1. A couple of hand augers were performed at Landslide FID 4 to determine the soil composition. Cutting and filling activities were undertaken at the other three landslide areas for the

Table 1 Landslides where field reconnaissance was performed

FID	Coordinates at landslide centroid	Area of landslide (m^2)	Formation
1	47°48′05.1″ N, 102°37′46.6″ W	104,641	Sentinel Butte
2	47°40′38.4″ N, 102°54′18.8″ W	2,647,300	Sentinel Butte
3	47°31′42.1″ N, 102°43′04.8″ W	269,387	Sentinel Butte
4	47°33′35.4″ N, 103°05′46.5″ W	2,071,500	Sentinel Butte

construction of well pads and roads, allowing visual inspection to be performed to determine the soil composition at those locations. Visual inspection of landslides located along Highway 22 of Fort Berthold, North Dakota, and 108th Avenue Northwest Watford City, McKenzie County, North Dakota was performed from a car. Landslide FID 1 was located near the bank of the Missouri River, whereas the other three landslides were not located near any water source, as verified from Google Earth. Site reconnaissance for four landslides and visual inspection of several other landslides revealed exposed dissolved salts on the failure surfaces, with large quantities of salt being leached by rainfall and accumulating at the toe of the landslides, as shown in Fig. 5.

Changes in the pore fluid salinity will affect the strength of the slope materials and, thus, their stability (Tiwari and Ajmera 2015). The distribution of landslides with salt type and concentration is important given the ubiquitous presence of dissolved sulfates in the pore fluids of the study area. Data related to the EC of the soil and the SAR were obtained from the Soil Survey Geographic Database (SSURGO, USDA 2020). SSURGO uses the saturated paste extract method to determine the EC and SAR. In particular, soil is saturated with water, which is then extracted under a partial vacuum to determine dissolved salts. This technique is described in USDA (2013) and Corwin and Yemoto (2017), and the details are not repeated here due to space limitations. The raw EC values from the data were converted into total soluble salts (TSS; mmol/L) and total dissolved solids (TDS; mg/L) using equations from Corwin and Yemoto (2017). SAR represents the ratio of sodium ion concentration ($[Na^+]$) to the concentrations of calcium ($[Ca^{2+}]$) and magnesium ($[Mg^{2+}]$) ions, as shown in Eq. (1).

$$\mathrm{SAR} = \frac{[Na^+]}{\sqrt{\frac{[Ca^{2+}]+[Mg^{2+}]}{2}}} \tag{1}$$

The general composition of the pore fluids in the soil found in North Dakota typically has three separate salt compounds: sodium sulfate (Na_2SO_4), calcium sulfate ($CaSO_4$), and magnesium sulfate ($MgSO_4$). This is based on the work of Keller et al. (1986) and Derby et al. (2014), who studied the surface geology and soils in the state. The TSS, TDS, and SAR values were used to calculate the mass concentrations of the ions within the salts: $[Na^+]$, $[Ca^{2+}]$, $[Mg^{2+}]$, and $[SO_4^{2-}]$ (concentration of sulfate ions). This conversion accounted for 95% of the EC values without reaching the solubility limits of Na_2SO_4, $CaSO_4$, and $MgSO_4$. It was performed by inversely parameterizing the mass concentrations and minimizing the objective function via the generalized reduced gradient method (Ladson et al. 1978). The process was automated in Excel via a Visual Basic for Application Macro.

The maximum likelihood method was used to fit probability density functions (PDF) to the distributions of frequency for various factors. In particular, this method identifies the parameters that will maximize the likelihood of observing the selected distribution under the assumption that each data point is statistically independent. Equation (2) describes the probability density function of a normal distribution, with Eqs. (3) and (4) representing the maximum likelihood method parameters. $f(x)$ represents the likelihood of the occurrence of x, the independent factor, σ is the standard deviation, μ is the mean, and n is the number of locations. If the distribution is a log-normal, Eq. (5) represents the PDF, while Eqs. (6) and (7) were used to determine the associated parameters. Lastly, exponential distributions

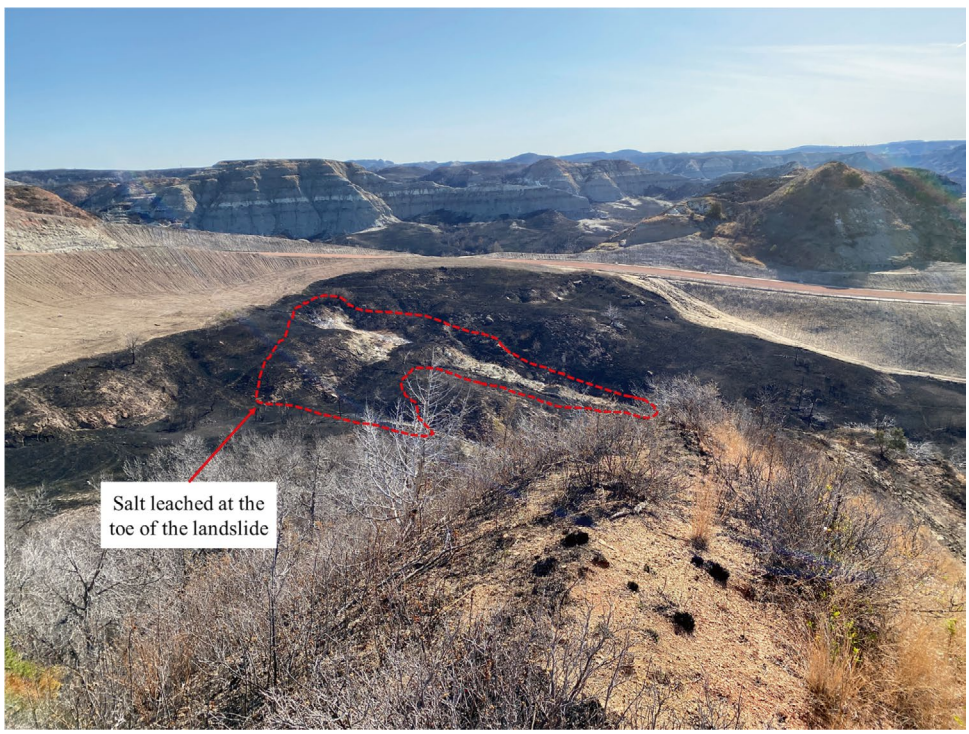


Fig. 5 Example of leached salt at landslide toe

are described by Eqs. (8) and (9). In Eq. (9), μ is calculated using Eq. (4). To measure the goodness-of-fit for each distribution, the Kolmogorov–Smirnov (K-S) test was used. In this test, the maximum difference between the observed frequency data and the fitted PDF is calculated. This difference is, then, compared with the critical difference for significance levels of 1 and 5%. These significance levels indicate a 1 or 5% chance, respectively, that the observed data do not follow the fitted PDF function. If the maximum difference exceeds the critical difference, then the likelihood that the PDF function does not fit the observed data is higher than the significance level. The critical differences for 1 and 5% significances were calculated from Eqs. (10) and (11) in Table 2, respectively. The corresponding values are also shown in Table 2.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \tag{2}$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{3}$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4}$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}x} \exp\left[\frac{-(\ln(x) - \mu)^2}{2\sigma^2}\right] \tag{5}$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\ln x_i - \mu\right)^2 \tag{6}$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\ln x_i \tag{7}$$

$$f(x) = \lambda e^{-\lambda x} \tag{8}$$

$$\lambda = \frac{1}{\mu} \tag{9}$$

t tests were performed on the landslide and control point datasets for each factor, assuming unequal variances. Factors that have an α-value less than 0.05 are considered to be statistically significant. The α-value represents the significance level for the test and is directly related to the probability of rejecting the null hypothesis. Cohen's *d* value, calculated using Eq. (12), helps to understand how large the statistically significant difference is. Specifically, it represents the number of standard deviations between the means of the landslides and control point distributions (Goulet-Pelletier et al. 2018). In Eq. 12, the subscripts represent the landslide dataset (1) and the control point dataset (2). Cohen's *d* value is converted into an effect size for easier interpretation of the data. The effect sizes corresponding to different Cohen's *d* values are shown in Table 3, which is based on the recommendations from Sawilowsky (2009).

$$d = \frac{\mu_2 - \mu_1}{\sqrt{\frac{(n_1-1)\sigma_1^2 + (n_2-1)\sigma_2^2}{n_1+n_2-2}}} \tag{12}$$

Pearson's correlation matrix was developed for the correlation between each pair of conditioning factors. The data for all non-numerical factors such as geology, land cover, and counties were converted into ones and zeros. For example, if a landslide location had the Sentinel Butte Formation as its main surficial geologic formation, then a one is assigned for this formation, and all other formations are assigned a zero for that location. This results in a numerical value for each factor. Using the Excel Correlation Data Analysis Tool, each conditioning factor is linearly correlated with the other conditioning factors. For parameters, any correlation coefficient between 0.70 and 1.00 or − 0.70 and − 1.00 was considered to be strongly positive and strongly negative correlations, respectively. Similarly, correlation coefficients between 0.20 and 0.70 or − 0.20 and − 0.70 were considered moderate positive and negative correlations, respectively (Ratner 2009). All correlations between a factor and itself were ignored because it was always equal to unity, as was expected. Only moderate and strong correlations are shown in the paper due to space limitations.

Results for the distribution of landslides

A total of 66,894 landslides and 68,395 randomly generated control points are used to perform the analyses described in this paper. Figure 6 displays the distribution of the landslides and control points within the North Dakota County boundaries. The largest fraction of the slope failures occurred in McKenzie County (18,068 or 27.0% of the total), followed by Dunn (8309 or 12.4% of the total) and Billings (6391 or 9.6% of the total) counties. Out of the 53 counties that make up the state of North Dakota, Ramsey County was the only one with no slope failures. The top 10 counties accounted for 78% of the total number of landslides. For the control points, the top 3 counties were McKenzie (2594 or 3.8% of

Table 2 Critical differences for K-S test

Significance level	Equation for critical difference	Landslide locations ($n = 66{,}894$)	Control points ($n = 68{,}395$)
1%	$\frac{1.63}{\sqrt{n}}$ (10)	0.006	0.006
5%	$\frac{1.36}{\sqrt{n}}$ (11)	0.005	0.005

Table 3 Effect sizes and corresponding Cohen's *d* values (Sawilowsky 2009)

Cohen's *d* value/No. of Std. Dev	Effect size
0.1	Very small
0.2	Small
0.5	Medium
0.8	Large
1.2	Very large
2	Huge

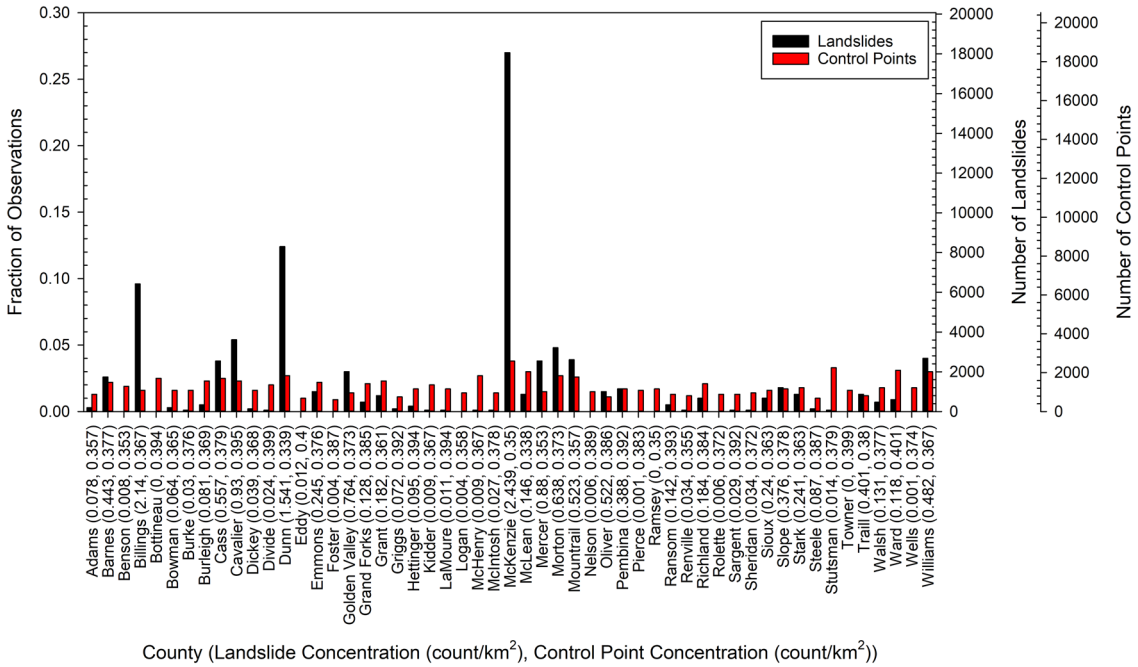


Fig. 6 Distribution of landslide and control points with North Dakota counties

the total), Stutsman (2258 or 3.3% of the total), and Ward (2138 or 3.1% of the total). Each county in North Dakota contained at least 130 control points.

The landslide concentration (LC) per county was determined as the total number of landslides per unit area of the county. The three counties with the highest concentrations of landslides are McKenzie (2.44 landslides/km^2), Billings (2.14 landslides per km^2), and Dunn (1.54 landslides per km^2). The landslide and control point concentrations are shown in the *x*-axis labels of Fig. 7. The three counties with the largest concentrations of control points are Ward, Eddy, and Divide, with each approximately containing 0.40 control points per km^2.

The areas of the landslides included in this study ranged from about 2 to 8,000,000 m^2. The exponential distribution of the area of these landslides has a mean of 19,797 m^2 and a standard deviation of 105,996 m^2, as shown in Fig. 7. The control point locations have no area as they represent randomly generated locations across the study area. The maximum difference in the K-S test was 0.28, which did not pass the critical values for both 1% (0.006) and 5% significance (0.005). However, this distribution is still considered to be appropriate for a number of reasons, as explained in the "Discussion" section of this paper.

The distribution of the landslide and control point locations in relation to surficial geology in North Dakota is shown in Figs. 1 and 8. The largest fraction of the landslides occurred in the Sentinel Butte Formation (34,063 or 51% of the total), followed by the Bullion Creek Formation (8695 or 13% of the total) and river sediments of the Oahe Formation (6421 or 10% of the total). As for the control point locations, the largest fraction was in gently undulating collapsed glacial sediments (10,766 or 16% of the total), followed by the Sentinel Butte Formation (8273 or 12% of the total) and in the undulating collapsed glacial sediments (6071 or 9% of total). The landslide concentration corresponds to the total number of landslides per unit area of the geologic formation and is shown in Fig. 8. As displayed in the figure, the formations with the highest concentrations of landslides are the Undivided Niobrara and Carlile Formation (11.93 landslides per km^2), Pierre Formation (3.20 landslides per km^2), and Undivided Upper and Middle Tertiary Rock (2.10 landslides per km^2). The highest control point concentrations were found in the eroded lake sediments of the Coleharbor Formation (0.73 control points per km^2), glacial sediments on subglacially molded surfaces (0.44 control points per km^2), and the Undivided Niobrara and Carlile Formations (0.43 control points per km^2).

The formations with the largest number of landslides and control points correspond to the formations with the largest areas. This might be due to the fact that the larger the area of a formation, the greater the likelihood that there will be a slope failure in that area. Since the control points are randomly generated across the entire study area, the larger formations will contain more points.

The same trend occurs for the geologic formations with the highest landslide concentration. The top three formations with the highest landslide concentrations encompass some of the smallest areas of the surface of North Dakota. The distributions for landslides and control points in relation to geologic formations show that the size of the geologic formation plays a large role in the probability of landslides occurring in that formation. In the *t* test results, all of the geologic formations had statistically significant differences between the landslide and control point locations (Table 4). Cohen's *d* value indicated a very small effect size for each formation except the Sentinel Butte Formation. It is noted that with a very large dataset (such as those used in this study), small differences may suggest statistically significant differences. These results indicate that surficial geology plays an important role in the conditioning of slopes for potential instabilities.

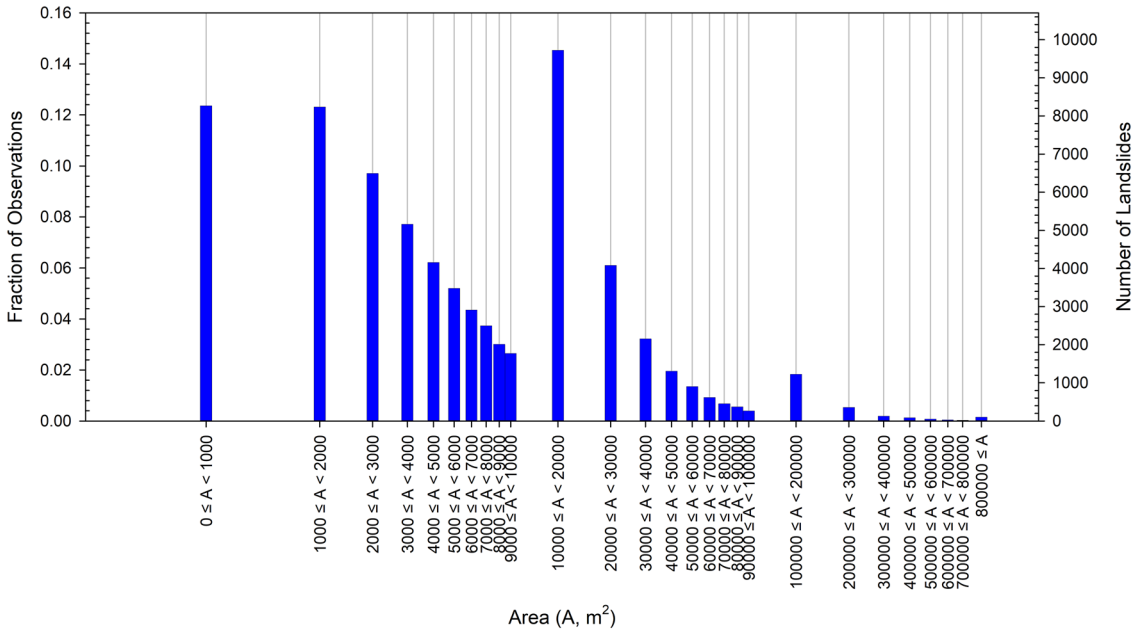


Fig. 7 Distribution of landslide areas

The largest effect size (small) was observed in the Sentinel Butte Formation. This formation is known to have poorly cemented angular-grained sandstone with high clay contents, along with poorly cemented claystone and limestone. There is also evidence of expanding clays within the Sentinel Butte Formation. The combination of these varying weak soils and sedimentary rocks leads to cross-stratification, which is prone to sliding failures (Wallick 1984).

The topography of North Dakota varies dramatically from flat grasslands to nearly vertical rock faces, as illustrated in Fig. 4. Figure 9 shows that landslides were typically found in areas with slopes less than 40°, with the majority of the failures occurring on slopes with inclinations between 9° and 14°. The distribution for the slope inclination of the landslide locations followed a lognormal distribution that had a mean of 13.4° and a standard deviation of 6.8°. The maximum difference in the

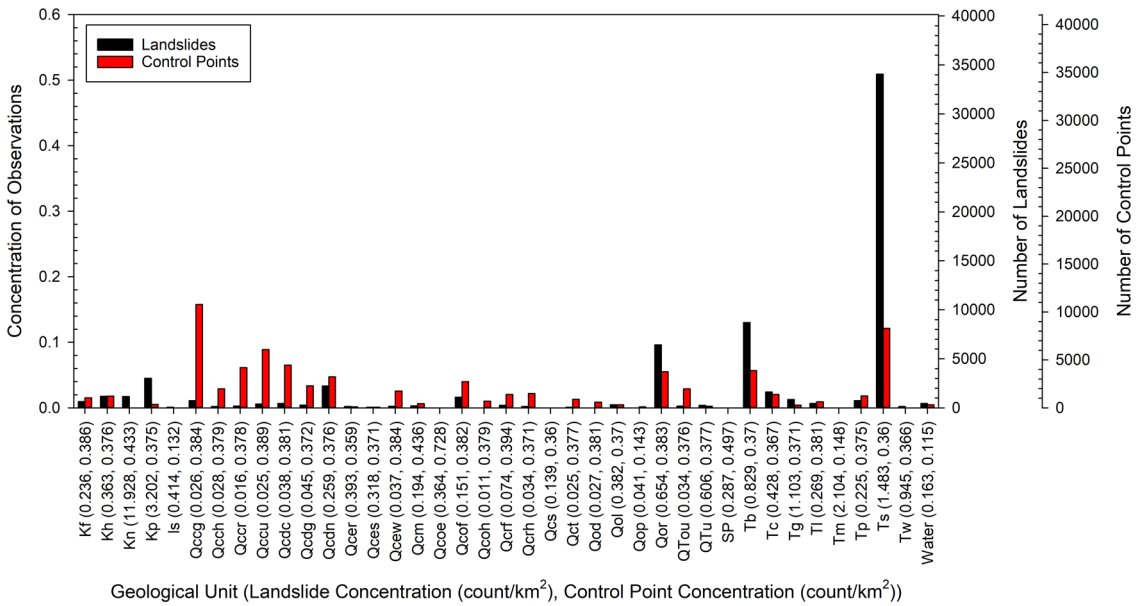


Fig. 8 Distributions of landslide and control points with geologic formations in North Dakota. The formation names associated with each geologic unit are summarized in Fig. 1

Table 4 Statistically significant *t* test and Cohen's *d* value results for geology

Formation	*p* value	Cohen's *d* value	
Kf	< 0.001	− 0.003	Very small effect size
Kh	< 0.001	0.000	Very small effect size
Kn	< 0.001	0.008	Very small effect size
Kp	< 0.001	0.019	Very small effect size
Is	< 0.001	0.000	Very small effect size
Qccg	< 0.001	− 0.068	Very small effect size
Qcch	< 0.001	− 0.012	Very small effect size
Qccr	< 0.001	− 0.027	Very small effect size
Qccu	< 0.001	− 0.039	Very small effect size
Qcdc	< 0.001	− 0.027	Very small effect size
Qcdg	< 0.001	− 0.014	Very small effect size
Qcdn	< 0.001	− 0.006	Very small effect size
Qcer	< 0.001	0.000	Very small effect size
Qces	< 0.001	0.000	Very small effect size
Qcew	< 0.001	− 0.011	Very small effect size
Qcm	< 0.001	− 0.002	Very small effect size
Qcoe	< 0.001	0.000	Very small effect size
Qcof	< 0.001	− 0.011	Very small effect size
Qcoh	< 0.001	− 0.005	Very small effect size
Qcrf	< 0.001	− 0.008	Very small effect size
Qcrh	< 0.001	− 0.009	Very small effect size
Qcs	< 0.001	0.000	Very small effect size
Qct	< 0.001	− 0.006	Very small effect size
Qod	< 0.001	− 0.004	Very small effect size
Qol	< 0.001	0.000	Very small effect size
Qop	< 0.001	0.000	Very small effect size
Qor	< 0.001	0.019	Very small effect size
QTou	< 0.001	− 0.012	Very small effect size
QTu	< 0.001	0.001	Very small effect size
SP	< 0.001	0.000	Very small effect size
Tb	< 0.001	0.034	Very small effect size
Tc	< 0.001	0.002	Very small effect size
Tg	< 0.001	0.004	Very small effect size
Tm	< 0.001	0.000	Very small effect size
Tp	< 0.001	− 0.003	Very small effect size
Ts	< 0.001	0.181	Small effect size
Tw	< 0.001	0.001	Very small effect size

The formation names are provided in Fig. 1

K-S test was 0.107, which is greater than the critical value for both 1% (0.006) and 5% significance (0.005). For the control points, the data followed an exponential distribution, as shown in Fig. 9, with 87% of the control points falling in regions with slope inclinations between 0° and 4°. The distribution had a standard deviation of 3.1° and a mean of 2.0°. The K-S values for the control point locations had a maximum difference of 0.111, which was greater than the critical values for both 1% (0.006) and 5% significance (0.005).

The *t* test indicated a statistically significant difference between the landslides and control points with a *p* value less than 0.001. The Cohen's *d* value for slope inclination was equal to 5.3, which represents a huge effect size. The statistically significant differences in the *t* test indicate that the differences in the slope inclinations between the landslide and control point locations are not an artifact of random chance. Rather, slope inclination is a critical factor that significantly differentiates between locations with landslides and those without. This agrees with the widely accepted fact that in slope stability, the slope inclination of the face of the soil mass will have a large impact on its stability.

Godt et al. (2012) presented a straight-line relationship between the topographic slope and the relief in landslide locations based on landslide inventories in New Jersey, the San Francisco Bay region, Oregon, New Mexico, and North Carolina. The relief in North Dakota is stated to be between 91 and 152 m in the North Dakota Badlands and the Missouri Coteau (Enz 2003), both of which are primarily located in the western part of the state. The Missouri Coteau is approximately a 48- to 113-km-wide section that starts in the northwest corner of the state, diagonally extending to the central part of the southern border (Enz 2003). These physiographic regions typically have higher slope inclinations, as shown in Fig. 4. Glaciated Plains and the Red River Valley encompass the majority of the remainder of the state and have relatively low relief. In fact, Enz (2003) suggested that the relief in the Glaciated Plains was typically less than 8 m, reaching a maximum of 30 m, while the Red River Valley is essentially flat. Thus, the fact that the majority of the slopes fail on slopes with inclinations between 9° and 14° appears to corroborate the relationship between slope inclination and relief proposed by Godt et al. (2012). It is noted that Godt et al. (2012) used Shuttle Radar Topography Mission (SRTM) elevation data with an approximately 90 km resolution and an approximately 1 km moving window to calculate the relief. Enz (2003) does not provide any details regarding the data and process used to calculate the relief but rather provides some absolute relief values. Thus, a direct comparison of the relief values from Enz (2003) to the relationship suggested by Godt et al. (2012) should be made cautiously. However, the absolute values from Enz (2003) suggest that the relationship proposed is valid in western North Dakota, although it does not capture the mapped failures in eastern North Dakota. This suggests that the relationship may need refinement using additional data from low relief areas (such as those found in eastern North Dakota), better resolution for the slope and/or relief data, or the incorporation of other factors, such as geology, land use, and precipitation, which were not considered by Godt et al. (2012).

The distribution of landslides and control points in relation to the land cover use in North Dakota is shown in Fig. 10. In this study, 29% of the landslides occurred in areas with herbaceous vegetation, which

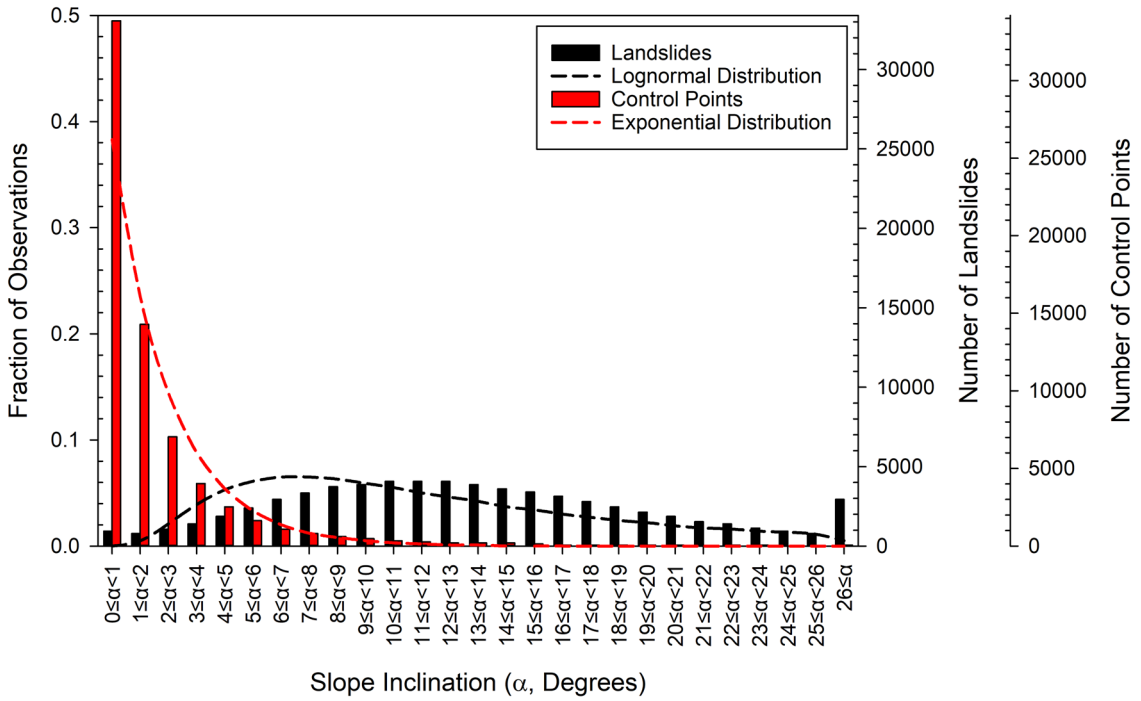


Fig. 9 Distribution of landslide and control points with slope inclination

is defined as grassland areas with plants classified as graminoids. The graminoids in North Dakota include the prairie June grass, western wheatgrass, big bluestem, little bluestem, Indiangrass, blue grama, and green needlegrass. The second largest land cover encountered in the landslide locations is shrubs/scrubs, which account for 25% of the total failures. Shrubs/scrubs include indigo bushes, lead plants, young trees (less than 5 m in height), and running serviceberries. With about 23% of the landslides, deciduous forests are the land cover with the third largest fraction of landslides. The top 3 land cover classifications for the control point locations were cultivated crops (53%), herbaceous vegetation (23%), and hay/pasture (6%).

The LC was also calculated for each land cover. The LC value for each classification is shown in the horizontal axis in Fig. 10. The three land covers with the highest LC were mixed forests (13.76 landslides per km^2), evergreen forests (9.47 landslides per km^2), and deciduous forests (6.11 landslides per km^2). High-intensity development (0.48 control points per km^2), low-intensity development (0.41 control points per km^2), and hay/pasture lands (0.39 control points per km^2) land covers were the top 3 land covers among the control point locations.

The largest land cover use in North Dakota is cultivated crops, comprising 51% of the study area. Although 53% of control points fell in regions with this land cover, only 4% of the landslides occurred in them. This is expected since farmland is usually selected to be in relatively flat regions for easy harvesting. About 30–40% of the farmland in North America has drainage systems installed to help manage the surface and groundwater (Vlotman et al. 2020). The second largest land cover in the study area is herbaceous, which, as described earlier, includes seasonal grasses and bushes with weak root systems. Herbaceous plant life can have trouble growing deep and dense root systems when there are high silt and clay contents. Given the low density of root structures in the soil, erosion and instability of the topsoil is more likely (Dunaway et al. 1994). The topsoil in the study area has high clay contents, as discussed previously. All land cover uses in the study area were found to have statistically significant differences between the landslides and the control points based on the *t* test results (Table 5). Of those, only two had a Cohen's *d* value indicating a small statistical difference between the landslide and control point locations. These two classifications were cultivated crops and deciduous forests. These results indicate that there are differences in land cover use between the landslide and control point locations that are more than a random occurrence. In other words, land cover use plays a critical role in instability occurrences within the study area.

Figure 11 shows the distribution of the landslide locations and control points in relation to average annual rainfall in North Dakota. The rainfall data presented in Fig. 11 are the average from 2014 to 2019. The distribution of landslides with the amount of rainfall follows a lognormal distribution. The mean and standard deviation of the distribution are equal to 70.6 and 7.5 cm/year, respectively. The K-S values for the landslides found a maximum difference of 0.260, which was greater than the critical value for both 1% (0.006) and 5% significance (0.005). The control points followed a normal distribution. The mean, as seen in Fig. 11, is at 65.8 cm/year. The standard deviation was equal to 7.4 cm/year. The K-S test for the control point distribution derived a maximum difference of 0.032, which was greater than the critical value for both 1% (0.006) and 5% significance (0.005).

Figure 12 displays the distribution of the landslides and control points in relation to the average annual snowfall in North Dakota. The snowfall shown in the figure is an average from 2010 to 2019. Both the landslide locations and control points followed normal

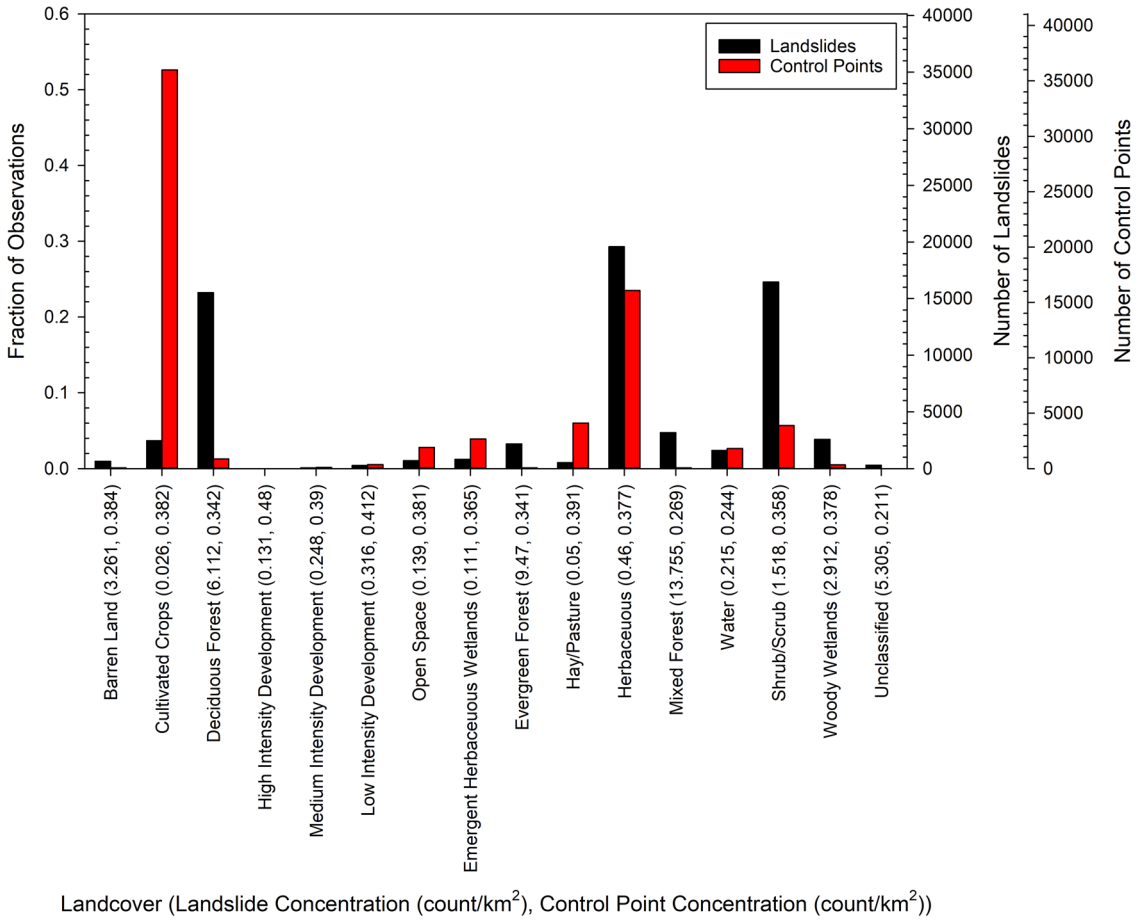


Fig. 10 Distribution of landslide and control points with land cover

distributions, with the majority of the locations corresponding to regions with 100 to 160 cm of snowfall per year. The mean for the landslide snowfall normal distribution was 130.0 cm/year, while the distribution for the control points had a mean of 122.5 cm/year. The K-S test results indicated a maximum difference for the landslides (0.258) larger than the critical values for 1% and 5% significance. The distribution for the control points had a maximum difference of 0.039, which is greater than the critical values for 1% (0.006) and 5% (0.005) significance. Figures 11 and 12 should be used with caution due to the fact that the data are based on the averages of multiple years of snowfall and rainfall data. This average was utilized because the exact dates of failure for the landslides are not known. Although there is some variability, the data are still valuable for examining the trends.

A statistically significant difference (p value < 0.001) was observed between the rainfall at the landslide locations and the control points. Cohen's d value was found to be equal to -0.221, which correlates to a small effect size. Similarly, a statistically significant difference (p value $= 0.003$) with a Cohen's d value of 0.110 (small effect size) was observed between the snowfall distributions. The data indicate that landslides were more likely to occur in areas that received larger amounts of rainfall and snowfall. At rainfalls greater than 80 cm/year or snowfalls greater than 180 cm/year, the data show lower probabilities of slope failure. This is interesting since saturated soils are generally weaker than dry or partially saturated soils. Furthermore, water from precipitation events or infiltration from snow melt will increase the weight of the soil, making it more susceptible to failure. The observations in this study may be an artifact of averaging the rainfall and snowfall data over several years. As a result, if these landslides occurred in the wetter years, they would not be captured by the averages used.

The distribution of landslides and control points with SAR is shown in Fig. 13. Those locations where the SAR could not be determined based on the data from the SSURGO database were excluded from this part of the analysis. The distributions for landslides and control points follow an exponential profile with means of 1.36 and 1.72 and standard deviations of 1.62 and 2.56, respectively. Over 87% of landslide locations and control points had a SAR of less than 3.5. The exponential distributions did not satisfy the K-S test, having a maximum difference of 0.041 and 0.129 for the landslides and control point distributions, respectively. SAR was determined to have a statistically significant difference between the landslide locations and control points with a p value less than 0.001. The Cohen's d value test determined that the difference was effectively small (-0.15).

Figure 14 shows the distribution of landslides and control points with EC. Those locations where the EC could not be determined based on the data from the SSURGO database were

Table 5 Statistically significant *t* test and Cohen's *d* value results for land cover

Land cover	*p*-value	Cohen's *d* value	
Barren land	< 0.001	0.004	Very small effect size
Cultivated crops	< 0.001	−0.229	Small effect size
Deciduous forest	< 0.001	0.103	Small effect size
Developed, high intensity	< 0.001	0.000	Very small effect size
Developed, low intensity	< 0.001	−0.001	Very small effect size
Developed, medium intensity	< 0.001	0.000	Very small effect size
Developed, open space	< 0.001	−0.008	Very small effect size
Emergent herbaceous wetlands	< 0.001	−0.013	Very small effect size
Evergreen forest	< 0.001	0.015	Very small effect size
Hay/pasture	< 0.001	−0.024	Very small effect size
Herbaceous	< 0.001	0.027	Very small effect size
Mixed forest	< 0.001	0.022	Very small effect size
Open water	< 0.001	−0.001	Very small effect size
Shrub/scrub	< 0.001	0.088	Very small effect size
Woody wetlands	< 0.001	0.016	Very small effect size

excluded from this part of the analysis. The landslide data follow a lognormal distribution, with the majority of the locations having EC values between 0 and 5.5 dS/m. The mean of the distribution is 1.69 dS/m. The K-S test for the landslide locations had a maximum difference of 0.107, which was greater than the critical value for both 1% (0.006) and 5% significance (0.005). The control points also follow an exponential distribution, with 73% of the locations having EC values ranging from 0 to 2.5 dS/m. With a

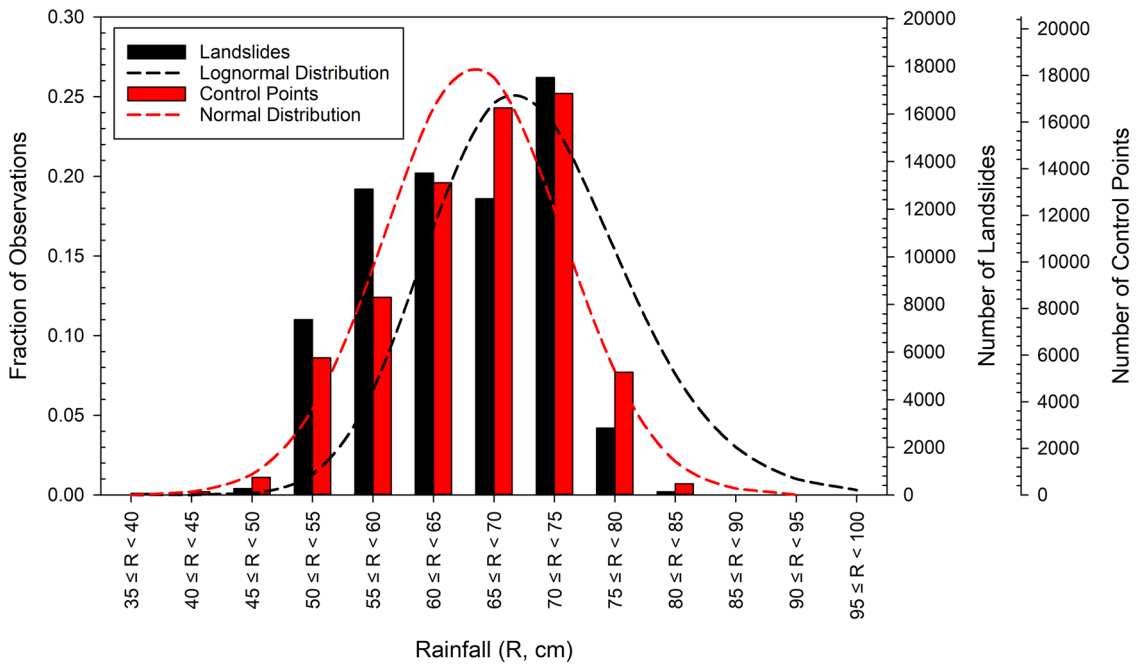


Fig. 11 Distribution of landslide and control points with average rainfall from 2014 to 2019

maximum difference of 0.175, the control points also did not satisfy the K-S test for both 1% (0.006) and 5% significance (0.005).

It is generally accepted that soils with an EC less than 4 dS/m and a SAR less than 13 are classified as non-saline and non-sodic soils (Davis et al. 2012). Thus, soils at the landslide locations and the control point can be classified as non-saline and non-sodic soils. Sodicity, which refers to the presence of high levels of exchangeable sodium ions (Na^+) in the soil, can have a detrimental effect on soil structure. This is because Na^+ ions have a larger hydrated radius compared to other cations such as calcium (Ca^{2+}) and magnesium (Mg^{2+}), which are normally present in soil exchange sites in the study area (Davis et al. 2012). The displacement of Ca^{2+} and Mg^{2+} by Na^+ ions can lead to a breakdown of soil aggregates and cause soil particles to disperse, resulting in a degradation of the soil structure. The dispersal of soil particles can result in soil crusting, surface sealing, increased soil erosion, reduced water infiltration, and decreased plant growth (Davis et al. 2012). All of these factors can increase the likelihood of stability issues.

As discussed earlier, EC and SAR were used to calculate the salt concentrations in the soil pore fluids. For the landslide locations, the calcium, magnesium, and sodium sulfate concentrations varied greatly, with ranges of 0 to 2870 mg/L, 0 to 4639 mg/L, and 0 to 5270 mg/L, respectively. These ranges for the control points were 0 to 4689 mg/L, 0 to 10,455 mg/L, and 0 to 7841 mg/L, respectively. Figures 15, 16, and 17 show the distribution of the landslides and the control points with calcium, magnesium, and sodium sulfate concentrations, respectively. About 98% of the landslide locations and control points had salt concentrations below 1100 mg/L. The calcium sulfate concentrations followed a lognormal distribution for both the landslide locations and control points, with a mean of 306.4 and 384.0 mg/L, respectively. The magnesium sulfate concentration for the landslide locations and control points also followed a normal distribution with a mean of 424.1 and 501.6 mg/L, respectively. The sodium sulfate concentrations followed a lognormal distribution for both the landslide locations and control points. The datasets have means of 270.2 and 407.4 mg/L for the landslide locations and control points, respectively. The maximum differences for all the salt concentrations from both the landslide locations and control points were above the critical values in the K-S test.

Statistically significant differences were observed between the calcium sulfate, magnesium sulfate, and sodium sulfate distributions. The Cohen's *d* values for all three indicated very small effect sizes. High sodium concentrations in clayey soils can result in the dispersion of the particles as the electrostatic bonds between the particles weaken when the sodium concentration increases (Pearson 2020). The calcium and magnesium in high concentrations in the soil cause the clayey soil to form a more flocculated structure (Pearson 2020), but this is dependent on their ratio to the sodium concentration in the soil. If the sodium concentration is higher, it will react with the soil, resulting in a dispersed structure. Thus, the lower mean concentrations in the landslide locations are leading to a more dispersed clay structure, which is in agreement with the SAR and EC results. Low salt concentrations in the pore fluid of clay-rich soil can be indicative of soil with thick, diffused double layers. This can lead to weak intermolecular bonding between the clay particles (Mitchell and Soga 2005), which can be related to lower shear strengths leading to instabilities.

Discussion

In the majority of cases, the maximum difference from the K-S tests for PDFs exceeded the critical difference. However, the provided distributions are still believed to be the best representations of the data for a number of reasons. First, the maximum likelihood method used to estimate the parameters assumed each data

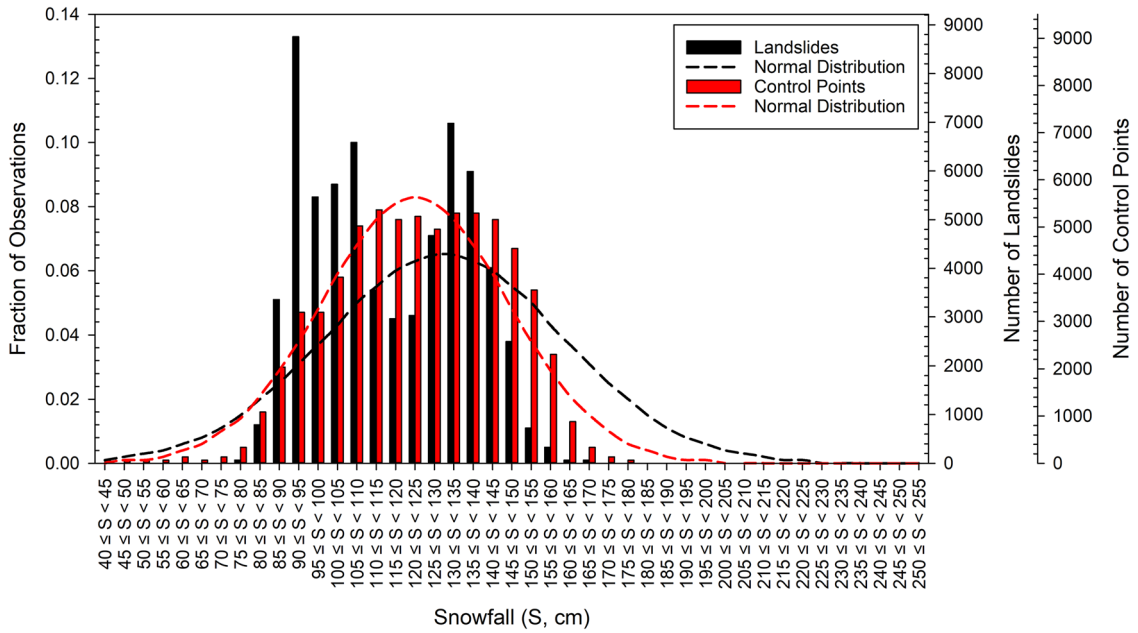


Fig. 12 Distribution of landslide and control points with average snowfall from 2010 to 2019

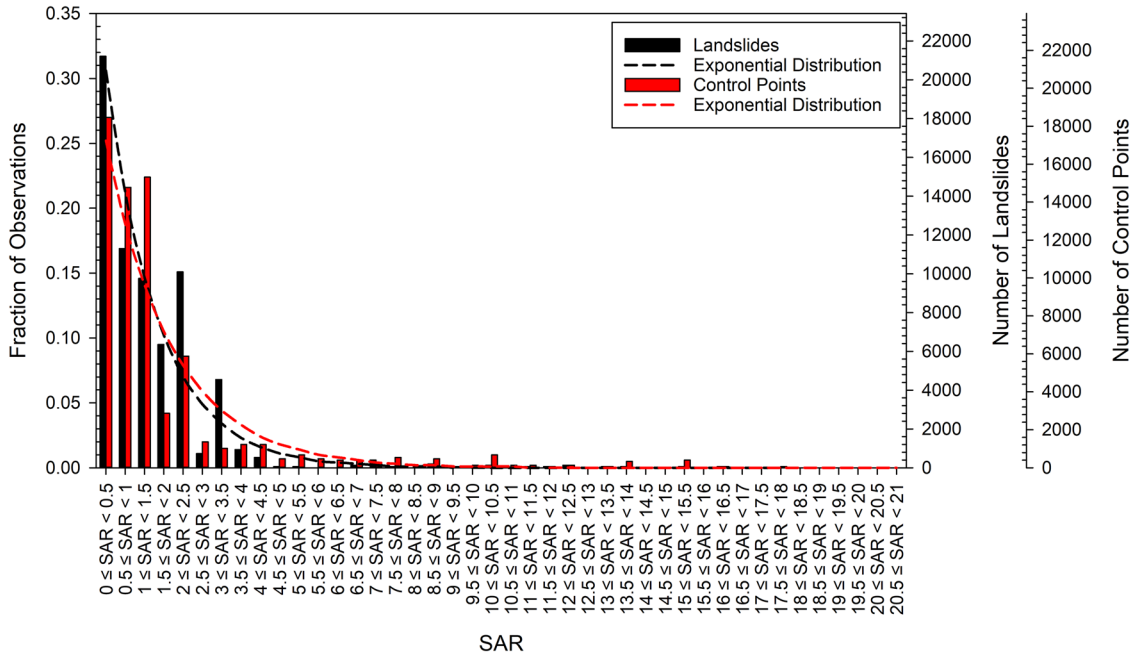


Fig. 13 Distribution of landslide and control points with sodium absorption ratio (SAR)

point was statistically independent. However, landslides can be correlated or can even cause other landslides. Thus, the PDFs developed should be utilized with caution. Second, a single value was assigned for each parameter at each landslide location. This may have resulted in an unintentional oversampling of a certain bin in the histograms developed, thereby skewing the maximum difference values. Third, in the K-S tests, the likelihood of rejecting the proposed theoretical PDF increases with an increase in the number of data points since small differences can be statistically significant in such datasets. Finally, the PDFs were selected to fit the physical meaning of the data. Thus, even though the critical difference is exceeded, the PDFs presented in this paper best represent the underlying physical phenomena.

Radbruch-Hall et al. (1982) identified two major regions in North Dakota as having moderate susceptibility to landslide occurrence. Both of these regions were stated to have low incidence levels. One of these

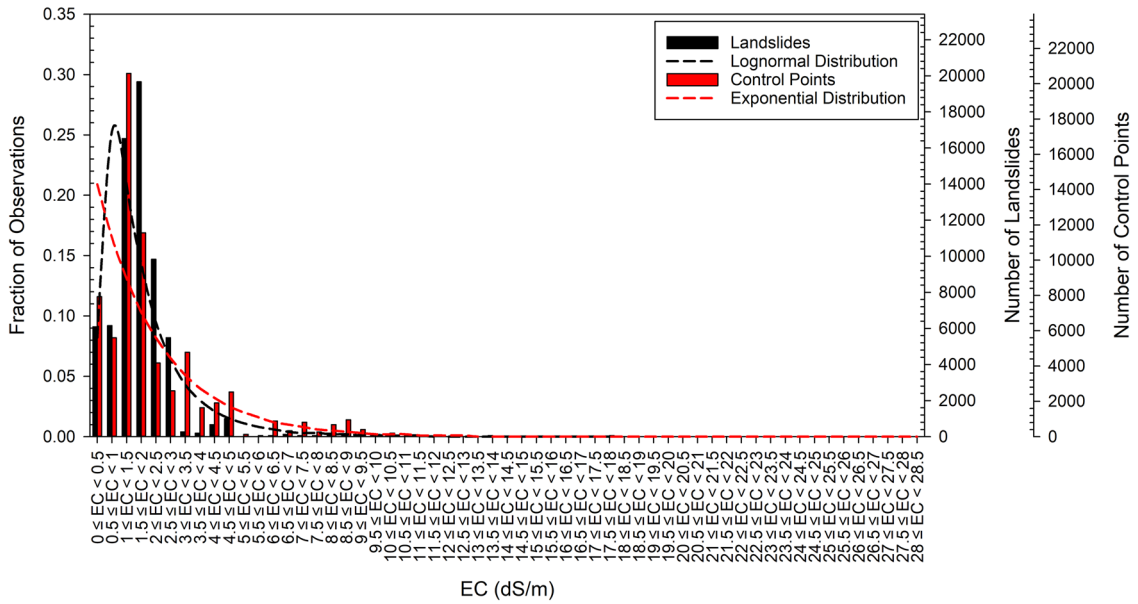


Fig. 14 Distribution of landslide and control points with electrical conductivity (EC)

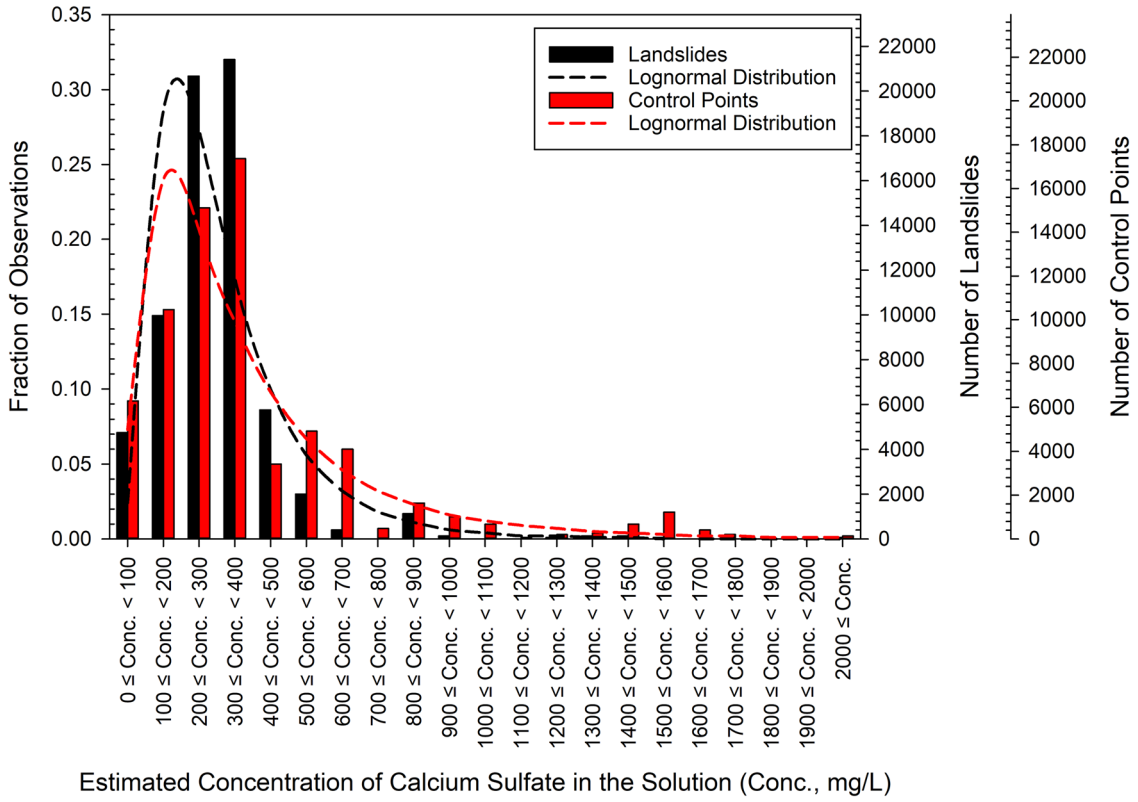


Fig. 15 Distribution of landslide and control points with calcium sulfate concentration at saturation

regions was the southwestern portion of the state, extending past the Missouri River and Lake Sakakawea. The other was along the central portion of the northern border of the state. Interestingly, many of the mapped landslides (Fig. 4) appear to be in the southwestern portion of the state that was identified by Radbruch-Hall et al. (1982) as having moderate susceptibility with low incidence. However, the majority of the mapped landslides appear to be located along the northern boundary of the region that was identified by Radbruch-Hall et al. (1982). On the other hand, only a handful of landslides were mapped in the second region identified as being moderately susceptible to landslides by Radbruch-Hall et al. (1982).

Interestingly, all of the landslides mapped in this study in the eastern part of the state (Fig. 4) fall into the regions classified by Radbruch-Hall et al. (1982) as having both low susceptibility and low incidence to landsliding. However, Godt et al. (2012) noted that while locations with negligible landslide susceptibility increased significantly east of longitude −105°, landslide susceptibility in this region was predominantly along river valley slopes (particularly those that drain to the east) and isolated upland regions. Although the aspect of the slope failures was not evaluated in this study, many of the mapped slope failures (Fig. 4) in the eastern part of the state are found along the Red River Valley.

In order to examine relationships between the conditioning factors, a correlation matrix was developed. A correlation matrix is a way of analyzing the linear relationships between multiple factors. For each pair of factors, the matrix summarizes a correlation coefficient indicating the goodness-of-fit and the type of relationship. Specifically, a positive value indicates an increase in the dependent variable as the independent variable increases. On the other hand, a negative value suggests that the dependent variable will decrease as the independent variable increases. Coefficients with magnitudes between 0.7 and 1 indicate strong correlations between the two factors, while those with magnitudes between 0.2 and 0.7 indicate a moderate correlation. For this study, only linear relationships were assumed for all factors. This assumption can result in lower correlation coefficients if the factors have a non-linear relationship. As explained before, when creating the correlation matrix, the data for all non-numerical factors such as geology and land cover were converted into ones and zeros. Selected results from the correlation matrixes for the landslide data and control point data are shown in Table 6. Due to limited space, only correlations that were considered moderate (± 0.2 to ± 0.7) or strong (± 0.7 to ± 1.0) for either the landslides or control points are shown.

As seen from the results, the Sentinel Butte, Pierre, and Bullion Creek Formations were the three formations with the largest number of landslides. The Sentinel Butte and Bullion Creek Formations comprise the majority of the surface geology in the western portion of North Dakota (Fig. 1). The portion of the state that encompasses these two formations also has more steeply inclined slopes (Fig. 4). Furthermore, these two formations are known for their weak sedimentary rocks, expansive clays, volcanic ash deposits, and

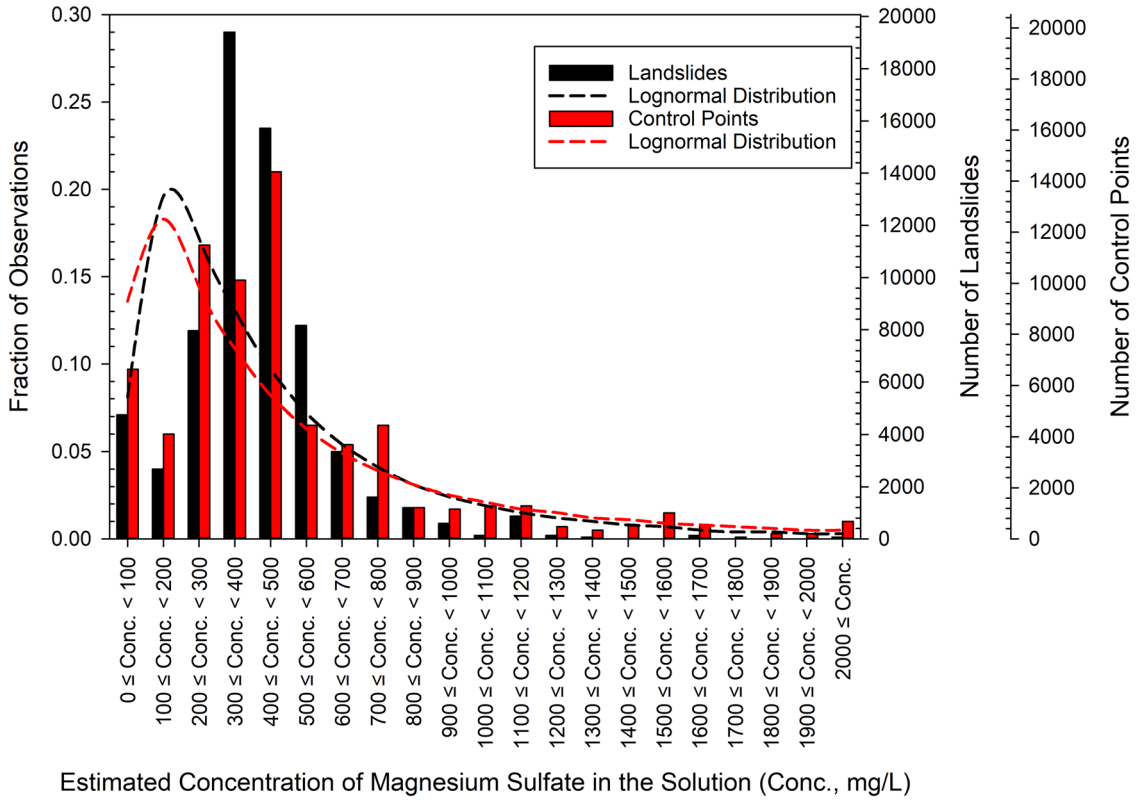


Fig. 16 Distribution of landslide and control points with magnesium sulfate concentration at saturation

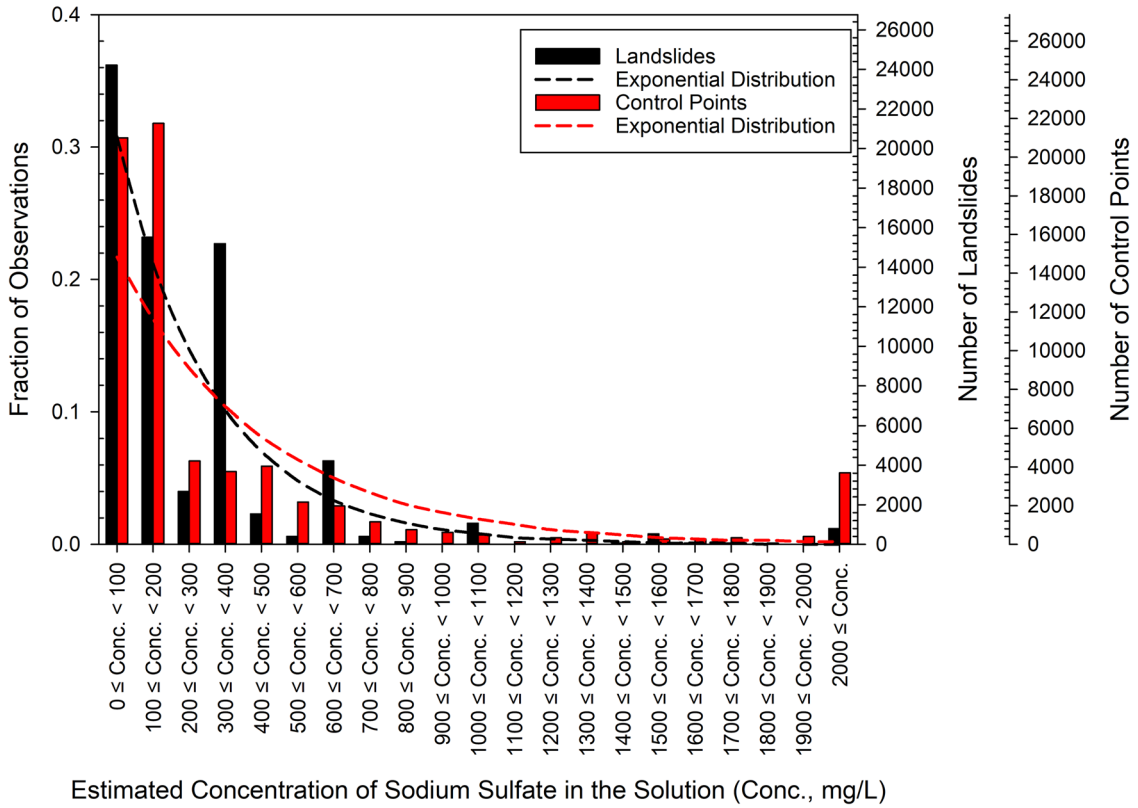


Fig. 17 Distribution of landslide and control points with sodium sulfate concentration at saturation

coal seams (Murphy et al. 2009). Thus, the three major geologic formations where landslides were most prevalent are known to have instability issues, suggesting that geology is a major conditioning factor in this region.

The highest landslide concentrations are seen in the Undivided Niobrara and Carlile Formation, Pierre Formation, and Undivided Upper and Middle Tertiary Rock. These formations are mainly made up of sedimentary rocks, including shale and limestone. (Kline 1942). Shale is known for its relatively high susceptibility to landslides due to its tendency to weather and erode easily when exposed to water and other environmental factors. The Upper and Middle Tertiary Rocks are susceptible to similar instabilities as the Sentinel Butte Formation and Bullion Creek Formation due to the similarities of their soil deposits. The Pierre Formation, on the other hand, makes up the majority of the eastern half of the state. Past studies have found that the materials found in the Pierre Formation tend to have low natural strengths, which will decrease with wetting or drying cycles (Schaefer and Bichmier 2013). Soils and sedimentary rock deposits in this formation were mainly composed of marine clays, which are highly susceptible to instability due to the leaching of salts in the soil.

The expansive clays in the Sentinel Butte and Bullion Creek Formations are susceptible to changes in salt concentrations. Depending on the concentration in the soil and the type of salt ($[Na^+]$, $[Ca^{2+}]$, or $[Mg^{2+}]$), the clay can either disperse or flocculate. If the soil disperses, it becomes denser with smaller pores, affecting plant growth. It also is weaker due to the parallel arrangement of the soil particles (Mitchell and Soga 2005), in which there are fewer surface-to-edge bonds. Such bonds are stronger than surface-to-surface bonds. The moderate correlation supports the observed notion that dispersion tends to occur in these landslide areas.

A small negative correlation (−0.12) between the Sentinel Butte Formation and rainfall was observed. The mean for the rainfall distribution for the landslides was lower than that of the control points, as seen in Fig. 11. Expansive clays have a tendency to crack during periods of low rainfall due to desiccation. Desiccation cracks can lead to the infiltration of water into the soil at faster rates, leading to the expansion of soils at greater depths while the surface is still dense. This phenomenon can lead to shear failure on slopes (Kodikara et al. 1999). Desiccation cracks, which are common in the semi-arid regions of the western USA, are also known to cause the phenomenon of piping in soil. Piping is where channels and tunnels are eroded under the surface of the soil. Desiccation cracks and piping occur in semi-arid regions that have rainfalls below 40 cm/year (Parker et al. 1990), which includes about 79% of the landslide locations. Piping has been reported to occur sometime in the Pliocene and Pleistocene eras in the White River and Sentinel Butte Formations in North Dakota (Parker et al. 1990). There is also evidence of piping occurring in the modern day due to the semi-arid climate and cold winters (Parker et al. 1990). Piping can cause landslides due to subsurface erosion or rapid saturation of the soil (Parker et al. 1990). On the other hand, natural processes such as the infiltration of rainfall can decrease the salt concentrations in the pore fluid of the soil. This will result in low EC values in the soil (Warrence et al. 2002), similar to the observations at the landslide locations. Figure 5 taken during the field reconnaissance show evidence of salt leaching occurring at landslide locations in the Sentinel Butte Formation. The decrease in salt concentrations has been directly correlated with a loss of shear strength in the soil (Tiwari and Ajmera 2015; Tiwari et al. 2005).

Table 6 Correlation matrix results for landslides and control points

Factor	Correlated to	Landslides	Control points
Average rainfall	EC	−0.22	−0.06
Average rainfall	SAR	−0.21	−0.07
Average rainfall	TDS	−0.22	−0.06
Average rainfall	$[Mg^{2+}]$	−0.25	−0.06
Average rainfall	$[SO_4^{2-}]$	−0.21	−0.07
Average snowfall	Average rainfall	0.45	0.27
$[Ca^{2+}]$	EC	0.96	0.98
$[Ca^{2+}]$	SAR	0.78	0.83
$[Ca^{2+}]$	TDS	0.96	0.98
$[Mg^{2+}]$	EC	0.78	0.77
$[Mg^{2+}]$	SAR	0.37	0.33
$[Mg^{2+}]$	TDS	0.78	0.77
$[Mg^{2+}]$	$[Ca^{2+}]$	0.78	0.75
$[Na^+]$	EC	0.87	0.90
$[Na^+]$	SAR	0.96	0.97
$[Na^+]$	TDS	0.87	0.90
$[Na^+]$	$[Ca^{2+}]$	0.86	0.91
$[Na^+]$	$[Mg^{2+}]$	0.45	0.45
SAR	EC	0.84	0.83
SAR	TDS	0.81	0.99
$[SO_4^{2-}]$	EC	0.98	0.99
$[SO_4^{2-}]$	SAR	0.83	0.97
$[SO_4^{2-}]$	TDS	0.98	0.99
$[SO_4^{2-}]$	$[Ca^{2+}]$	0.98	0.99
$[SO_4^{2-}]$	$[Mg^{2+}]$	0.79	0.78
$[SO_4^{2-}]$	$[Na^+]$	0.90	0.91
TDS	SAR	0.84	0.83
Qor	Cultivated crops	0.31	0.05
Qccg	Cultivated Crops	0.02	0.22
Shrub/scrub	Ts	0.21	0.33

This reduction in salt concentration with precipitation is noted in Table 6.

Many of the landslides in this study had land cover classifications of herbaceous and scrub/shrub (Fig. 10). The types of plant life related to this land cover are known to have shallow and weak root structures (ND 2021; NLCD 2016) or are seasonal plants (plants that die in the winter) with no wood-based structure. Strong deep root systems in soil can increase the shear strength of the soil and, therefore, provide more stability to the slope (Day 1993; Ghestem 2014; Zhang et al. 2010). However, the topsoil in these areas is not strongly stabilized by plant life. Furthermore, landslide locations with herbaceous and scrub/shrub tend to have lower recorded average annual rainfall (Table 6). As discussed earlier, lower annual rainfall in areas with clayey or silty soils leads to denser soils that prevent plants from growing deep, dense root systems. A low density of root structures in the soil will make erosion and instability of the topsoil more likely (Dunaway et al. 1994). Lowering the effective stabilization provided by the vegetation in herbaceous, scrub/shrub, and herbaceous wetland land covers.

Conclusions

A total of 66,894 landslides were mapped in North Dakota between the years of 1930 and 2023. A comprehensive analysis of the conditioning factors leading to slope failure was completed to discover which factors lead to an increased probability of failure in this region. Based on the results of that analysis, this study concluded the following six statements:

1. The largest fraction of the slope failures occurred in McKenzie County, followed in second and third place, respectively, by Dunn and Billings Counties. The three counties with the highest concentrations of landslides are also McKenzie, Billings, and Dunn.
2. All of the geological formations were found to be statistically significant from the *t* tests. However, a greater number of landslides were observed in the Sentinel Butte and Bullion Creek Formations. Failure in these formations is potentially due to expansive clays and weak sedimentary rock whose strength is susceptible to changes as a result of rainfall and fluctuations in the salt concentrations in the pore fluid.
3. The slope inclination of the landslide area was found to be statistically significant. The majority of the failures occurred on slopes with inclinations between 9° and 14°.
4. Many of the landslides in this study had land cover classifications of herbaceous and scrub/shrub. One possible reason for failures in these land covers may be related to the role of root reinforcement, but it requires further investigation in future studies.
5. According to the landslide distribution, the majority of the landslide locations experienced greater annual rainfall and snowfall than at the control points.
6. A majority of the landslide locations occurred in areas with low EC and SAR, which are indicative of regions with low salt concentrations in the pore fluid. Therefore, the landslide probability increases as the salt concentration decreases because a reduction in pore fluid salinity can cause a reduction in the shear strength of the soil.

Funding

The authors appreciate the financial support provided by the Geopier Foundation Company Midwest Scholarship, the Braun Intertec Corporation, the Department of Civil, Construction, and Environmental Engineering at Iowa State University, and the Department of Civil and Environmental Engineering (now, Civil, Construction, and Environmental Engineering) at North Dakota State University. Braun Intertac Corporation is also acknowledged for its support in conducting field investigations of the several landslide locations described in this study.

Data availability

Some or all data, models or codes that support the findings of this article are avialable from the corresponding author upon reasonable request.

Statements and declarations

The authors certify that they have no affiliations with or involvement in any organizations or entities with any financial or non-financial interest in the subject matter or materials discussed in this paper.

References

ACIS (2020) NOAA: regional climate centers. Appl Clim Info Sys.http://scacis.rcc-acis.org/

Anderson FJ, Maike CA (2017) Drones rising from the prairie: geological applications of unmanned aerial systems. GeoNews 44(2):8–14. https://www.dmr.nd.gov/ndgs/documents/newsletter/2017Summer/Drones%20Rising%20from%20the%20Prairie.pdf

Bluemle JP (1983) Geologic and topographic bedrock map of North Dakota, North Dakota geological survey. It can be found at: https://www.dmr.nd.gov/ndgs/documents/Publication_List/pdf/MisMaps/MM-25.pdf

Bluemle J, Biek B (2007) No ordinary plain-North Dakota's physiography and landforms. N Dakota Geol Surv. https://www.dmr.nd.gov/ndgs/ndnotes/ndn1.asp

City-Data (2020) North Dakota location, size and extent. North Dakota. http://www.city-data.com/states/North-Dakota-Location-size-and-extent.html

Clayton L (1980a) Geologic Map of North Dakota. US Geol Surv

Clayton L (1980b) Explanatory text to accompany the geologic map of North Dakota. North Dakota Geological Survey, Grand Forks

Corwin DL, Yemoto K (2017) Salinity: electrical conductivity and total dissolved solids. Soil Science Society of America 84(5):1442–1461

Davis JG, Waskom RM, Bauder JA (2012) Managing sodic soils. Colorado State University, Fort Collins, CO

Day RW (1993) Surficial slope failure: a case study. J Perform Constr Facil 7(4):264–269

Derby NE, Wick A, Casey FXM (2014) Spatially variable soil salinity, chemistry, and physical properties affect soil health. Soil Sci Soc Am Int Annual Meeting 2–5. Long Beach, CA

Dunaway D, Swanson SR, Wendel J, Clary W (1994) The effect of herbaceous plant communities and soil textures on particle erosion of alluvial streambanks. Geomorphology 9(1):47–56

EMDAT (2020) The International Disaster Database. Centr Res Epidemiol Disasters. https://www.emdat.be

Enz JW (2003) North Dakota topographic, climatic, and agricultural overview. North Dakota State University. https://www.ndsu.edu/fileadmin/ndsco/documents/ndclimate.pdf

Ghestem M, Veylon G, Bernard A, Vanel Q, Stokes A (2014) Influence of plant root system morphology and architectural traits on soil shear resistance. Plant Soil 377(1):43–61

Gill JR, Cobban WA (1965) Stratigraphy of the Pierre Shale, Valley City and Pembina Mountain Areas, North Dakota. US Government Printing Office, Washington

Godt JW, Coe JA, Baum RL, Highlighd LM, Keaton JR, Roth RJ (2012) "Prototype landslide hazard maps of the conterminous United States. Protecting Society through Improved Understanding, Landslides and Engineering Slopes, pp 245–250
Goulet-Pelletier J-C, Cousineau D (2018) A review of effect sizes and their confidence intervals, Part I: The Cohen's d family. The Quantitative Methods for Psychology 14(4):242–265
Highland L, Bobrowsky PT (2008) The landslide handbook: a guide to understanding landslides. US Geological Survey, Reston, Virginia
Jacob AF (1976) Geology of the Upper Part of the Fort Union Group (Paleocene), Williston Basin, with Reference to Uranium. North Dakota Geological Survey, Grand Forks
Keller LP, McCarthy GJ, Richardson JL (1986) Mineralogy and stability of soil evaporates in North Dakota. Soil Sci Soc Am J 50:1069–1071
Kline VH (1942) Stratigraphy of North Dakota. AAPG Bull 26(3):336–379
Kodikara J, Barbour SL, Fredlund DG (1999) Changes in clay structure and behaviour due to wetting and drying. In Proceedings 8th Australia New Zealand conference on geomechanics: consolidating knowledge (Vol. 179). Barton, ACT: Australian Geomechanics Society
Lasdon LS, Waren AD, Jain A, Ratner M (1978) Design and Testing of a Generalized Reduced Gradient Code for Nonlinear Programming. ACM Trans Math Softw 4(1):34–50
Mirus BB, Jones ES, Baum RL, Godt JW, Slaughter S, Crawford MM, Lancaster J, Stanley T, Kirschbaum DB, Burns WJ, Schmitt RG, Lindsey KO, McCoy KM (2020) Landslides across the USA: occurrence, susceptibility, and data limitations. Landslides 17:2271–2285. https://doi.org/10.1007/s10346-020-01424-4
Mitchell JK, Soga K (2005) Fundamentals of soil behavior, 3rd edn. John Wiley & Sons Inc, Wiley, New York
Moxness LD (2019) Twenty thousand slides and counting: recent advances in digital imagery expedite landslides mapping in North Dakota. GeoNews 17–19. https://www.dmr.nd.gov/ndgs/documents/newsletter/2019Winter/Twenty_Thousand_Slides_and_Counting.pdf
Moxness LD (2022) The first statewide landslide dataset: NDGS completes initial landslide mapping for North Dakota. GeoNews 12–15. https://www.dmr.nd.gov/ndgs/documents/newsletter/2022Winter/The_First_Statewide_Landslide_Dataset.pdf
Muñoz Sabater J (2019) ERA5-Land monthly averaged data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.68d2bb30
Murphy EC (2017) Landslides in North Dakota. GeoNews 44(1):1–5
Murphy EC, Nordeng SH, Juenker BJ, Hoganson JW (2009) North Dakota Stratigraphic Column: North Dakota Geological Survey Miscellaneous Series 91. N Dakota Geol Surv
ND (2020) Climate. Game and Fish. https://gf.nd.gov/wildlife/habitats/climate#:~:text=Annual%20precipitation%20ranges%20from%2013,to%204%20inches%20of%20rain
ND (2021) North Dakota plants and habitats overview. North Dakota Game and Fish Department. https://gf.nd.gov/wildlife/habitats/vegetation
NLCD (2016) National Land Cover Database 2016. Multi-Resolution Land Characteristics Consortium. https://www.mrlc.gov/data/legends/national-land-cover-database-2016-nlcd2016-legend#:~:text=Grassland%2FHerbaceous%2D%20areas%20dominated%20by,can%20be%20utilized%20for%20grazing
NOAA (2017) Climate of North Dakota. National Oceanic and Atmospheric Administration. https://web.archive.org/web/20070926010149/http://www5.ncdc.noaa.gov/climatenormals/clim60/states/Clim_ND_01.pdf
NOAA (2020) National Weather Service: national snowfall analysis. National Oceanic and Atmospheric Administration. https://www.nohrsc.noaa.gov/snowfall_v2/index.html?season=2009-2010&date=2011060912&version=3&format=.tif
Parker GG, Higgins CG, Wood WW (1990) Piping and pseudokarst in drylands. Geol Soc Am Spec Pap 252:77–110
Pearson KE (2020) Basics of salinity and sodicity effects on soil physical properties. Montana State University. https://waterquality.montana.edu/energy/cbm/background/soil-prop.html
Radbrunch-Hall DH, Colton RB, Davies WE, Luchhitta I, Skipp BA, Vanes DJ (1982) Landslide overview map of the conterminous United States. U.S. Geol Surv Prof Paper 1183
Ratner B (2009) The correlation coefficient: its values range between +1/− 1, or do they? J Target Meas Anal Mark 17(2):139–142
Ritchie H, Roser M (2014) Natural disasters. Our World in Data. https://ourworldindata.org/natural-disasters
Sawilowsky SS (2009) New effect size rules of thumb. J Mod Appl Stat Methods 8(2):597–599
Schaefer VR, Birchmier MA (2013) Mechanisms of strength loss during wetting and drying of Pierre Shale. Proc 18th Int Conf Soil Mech Geotech Eng 1183–1186
Schuster RL (1978) Introduction, Chapter 1. Trans Res Board Special Rep 176:1–10
Tiwari B, Ajmera B (2015) Reduction in fully softened shear strength of natural clays with NaCl leaching and its effect on slope stability. Journal of Geotechnical and Geoenvironmental Engineering 141(1):04014086
Tiwari B, Tuladhar GR, Marui H (2005) Variation in residual shear strength of the soil with the salinity of pore fluid. Journal of Geotechnical and Geoenvironmental Engineering 131(12):1445–1456
USDA (2013) National soil survey handbook. Title 430-VI
USDA (2020) Web soil survey. Nat Resour Conserv Serv. https://websoilsurvey.nrcs.usda.gov/
USGS (2016) NLCD land cover (CONUS). MRLC. https://www.mrlc.gov/data?f%5B0%5D=category%3ALand%20Cover
USGS (2019) USGS National Elevation Dataset (NED). ND State Water Commission. https://catalog.data.gov/dataset/usgs-national-elevation-dataset-ned
Vlotman WF, Smedema LK, Rycroft DW (2020) Modern land drainage: planning, design and management of agricultural drainage systems. CRC Press
Wallick BP (1984) Sedimentology of the bullion creek and sentinel butte formations (Paleocene) in a part of southern McKenzie County, North Dakota. Theses and Dissertations 311
Warrence NJ, Bauder JW, Pearson KE (2002) Basics of salinity and sodicity effects on soil physical properties. Department of Land Resources and Environmental Sciences, Montana State University-Bozeman, MT, 129:1–29
Zhang C-B, Chen L-H, Liu Y-P, Ji X-D, Liu X-P (2010) Triaxial compression test of soil–root composites to evaluate influence of roots on soil shear strength. Ecol Eng 36(1):19–26

Benjamin Shafer
Department of Civil, Construction, and Environmental Engineering, Iowa State University, Town Engineering, 813 Bissell Rd., Ames, IA 50011, USA
Email: brshafer@iastate.edu

Beena Ajmera (✉)
Department of Civil, Construction, and Environmental Engineering, Iowa State University, 490 Town Engineering, 813 Bissell Rd., Ames, IA 50011, USA
Email: bajmera@iastate.edu

Kamal Raj Upadhaya
Department of Civil, Construction, and Environmental Engineering, North Dakota State University, NDSU Dept. 2470, PO Box 6050, Fargo, ND 58108, USA
Email: kupadhaya@verdantas.com

Aaron Lee M. Daigh
Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Keim Hall (Keim) 369, Lincoln, NE 68583, USA
Email: adaigh2@unl.edu

Aaron Lee M. Daigh
Department of Biological Systems Engineering, University of Nebraska-Lincoln, Keim Hall (Keim) 369, Lincoln, NE 68583, USA
Email: adaigh2@unl.edu

Kamal Raj Upadhaya
Verdantas LLC, 5400 Limestone Rd., Wilmington, DE 19808, USA
Email: kupadhaya@verdantas.com